Mertz, P.
10/723,180
Sep 101

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:     September 8, 2005, 11:52:12 ; Search time 189.5 Seconds
                                          (without alignments)
                                          4702.353 Million cell updates/sec

Title:           US-10-723-180-1
Perfect score:   2181
Sequence:        1 gggcaggcaatgagagctg.........attactcagtctaaacaata 1152

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:      4211384

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

.Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10723180/runat_08092005_123255_29226/app_query.fasta_1.1351
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10723180_@CGN_1_1_224_@runat_08092005_123255_29226 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  A_Geneseq_16Dec04:*
         1: geneseqp1980s:*
         2: geneseqp1990s:*
         3: geneseqp2000s:*
         4: geneseqp2001s:*
         5: geneseqp2002s:*
         6: geneseqp2003as:*
         7: geneseqp2003bs:*
         8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1685	77.3	323	6	ABU56649	Abu56649 Lung canc
2	1685	77.3	323	7	ADN38936	Adn38936 Cancer/an
3	1685	77.3	323	5	ADS13763	Ads13763 Human aqu
4	1458.5	66.9	300	5	ABB57089	Abb57089 Mouse isc
5	617	28.3	265	5	ABG96280	Abg96280 Human ova
6	617	28.3	265	6	ABG97500	Abg97500 Novel hum
7	617	28.3	265	6	ABR92054	Abr92054 Human cer
8	617	28.3	265	7	ADD47950	Add47950 Human Pro
9	617	28.3	265	7	ADE62232	Ade62232 Human Pro
10	617	28.3	265	8	ADJ75631	Adj75631 Marker ge
11	617	28.3	265	8	ADJ75701	Adj75701 Marker ge
12	617	28.3	265	8	ADL70563	Adl70563 Cervical
13	614	28.2	265	2	AAW55787	Aaw55787 Rat aquap
14	614	28.2	265	2	AAW94319	Aaw94319 Rat aquap
15	614	28.2	265	7	ADD47948	Add47948 Rat Prote
16	614	28.2	265	8	ADE62230	Ade62230 Rat Prote
17	614	28.2	265	8	ADJ76346	Adj76346 Marker ge
18	614	28.2	265	8	ADJ76388	Adj76388 Marker ge
19	612	28.1	263	2	AAR25424	Aar25424 Human MIP
20	612	28.1	263	7	ADD48472	Add48472 Human Pro
21	612	28.1	263	7	ADE63588	Ade63588 Human Pro
22	603	27.6	118	3	AAG02008	Aag02008 Human sec
23	600.5	27.5	282	6	ABR41500	Abr41500 Human DIT
24	600	27.5	261	7	ADD48470	Add48470 Rat Prote
25	600	27.5	261	7	ADE63586	Ade63586 Rat Prote
26	597	27.4	271	8	ADS13759	Ads13759 Human aqu
27	594	27.2	271	7	AAR51070	Aar51070 A water c
28	589	27.0	271	7	ADG32000	Adg32000 Human hom
29	561	25.7	269	2	AAW55786	Aaw55786 Human aqu
30	561	25.7	269	2	AAW94322	Aaw94322 Human aqu
31	561	25.7	269	6	ABP58050	Abp58050 Human aqu
32	561	25.7	269	7	ADP65200	Adp65200 Human aqu
33	561	25.7	269	8	ABM80877	Abm80877 Tumour-as
34	561	25.7	269	8	ADS13757	Ads13757 Human aqu
35	525.5	24.1	222	8	ADL70567	Adl70567 Cervical
36	508	23.3	282	8	ADK50990	Adk50990 Human NOV
37	496.5	22.8	239	4	ABB64172	Abb64172 Drosophil
38	447	20.5	72	4	ABB29884	Abb29884 Peptide #
39	447	20.5	72	4	ABB20480	Abb20480 Protein #
40	447	20.5	72	4	AAM68252	Aam68252 Human bon
41	447	20.5	72	4	AAM55880	Aam55880 Human bra
42	447	20.5	72	5	ABG49906	Abg49906 Human liv
43	447	20.5	72	5	ABG37787	Abg37787 Human pep
44	443.5	20.3	253	3	AAG48889	Aag48889 Arabidops
45	443.5	20.3	278	3	AAG48888	Aag48888 Arabidops

ALIGNMENTS

```
RESULT 1
ABU56649
ID   ABU56649 standard; protein; 323 AA.
XX
AC   ABU56649;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polypeptide #242.
XX
KW   Lung cancer-associated polypeptide; cytostatic; emphysema;
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO200286443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US012476.
XX
PR   18-APR-2001; 2001US-0284770P.
PR   10-MAY-2001; 2001US-0290492P.
PR   09-NOV-2001; 2001US-0339245P.
PR   13-NOV-2001; 2001US-0350666P.
PR   29-NOV-2001; 2001US-0334370P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Aziz N, Murray R;
```

```
XX
DR   WPI; 2003-093161/08.
DR   N-PSDB; ABX76378.
XX
PT   Detecting a lung cancer-associated transcript in a cell from a patient
PT   for treating lung cancer, by contacting a biological sample from the
PT   patient with a polynucleotide that exhibits increased or decreased
PT   expression in lung cancer.
XX
PS   Claim 27; Page 375; 453pp; English.

CC   The invention relates to a method for detecting a lung cancer-associated
CC   transcript in a cell from a patient, comprising contacting a biological
CC   sample from the patient with a polynucleotide that selectively hybridises
CC   to a sequence that is at least 80 % identical to a gene that exhibits
CC   increased or decreased expression in lung cancer samples. Lung cancer-
CC   associated polynucleotides and polypeptides are used for identifying a
CC   compound that modulates a lung cancer-associated polypeptide, for
CC   inhibiting proliferation of a lung cancer-associated cell to treat lung
CC   cancer in a patient and for treating a mammal having lung cancer by
CC   administering a modulatory compound identified. The methods are useful
CC   for treating lung cancer, such as small cell lung cancer, non-small cell
CC   lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC   emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC   hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC   bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC   for diagnostic purposes and as targets for screening for therapeutic
CC   compounds that modulate lung cancer, such as antibodies. Sequences
CC   ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC   invention
XX
SQ   Sequence 323 AA;

Alignment Scores:
Pred. No.:                               Length:           323
Score:                 2.09e-158         Matches:          323
Percent Similarity:    1685.00           Conservative:       0
Best Local Similarity: 100.00%           Mismatches:         0
Query Match:           100.00%           Indels:             0
DB:                    77.26%            Gaps:               0
                       6

US-10-723-180-1 (1-1152) x ABU56649 (1-323)
```

```
Qy     40 ATGAGTGACGACGACCCACAGCCAAGGCGGTGGGTAAGTGTGGACCTTTGTGTACCAGAGAG 99
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetSerAspAspAspProThrAlaAspArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20

Qy    100 AACATCATGGTGGCTTCAAAGGGGTCTGGACTCAAAGCTTCTGGAAAGCAGTCACAGCG 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValAlaThrAla 40

Qy    160 GAATTTCTGCCATGCTTATTTTTGTTCTCCTCAGCCTGGATCCACCATCAACTGGGGT 219
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuSerThrIleAsnTrpGly 60

Qy    220 GGAACAGAAAAGCCTTACCGGTGACATGGTTCCATCTCCCTTGCTTTGGACTCAGC 279
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuGlyPheGlyLeuSer 80

Qy    280 ATTGCAACCATGGTGCAGTGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100

Qy    340 ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA 399
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleAla 120

Qy    400 GCCCAGTGCTGGGGCCATCATTGGAGCAGGAATCCTCTATCGGTCACACCTCCCAGT 459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer 140

Qy    460 GTGGTGGGAGGCCTGGGAGTCACCAGGTTCATCGGAAATCTTACCGGTCGTCATGGTCTC 519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    141 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160

Qy    520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC 579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 LeuValGluLeuIleIleThrPheAsnTrpCysTyrTyrLeuCysGlnLeuCysAspSer 180

Qy    580 AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTTCTGTTGCAATTGGA 639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleIleGlyPheSerValAlaIleGly 200

Qy    640 CATTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCT 699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220

Qy    700 GCAGTTATCATGGGGAAATTGGGAAAACCATTGGAATATATTGGGTTGGGCCCATCATAGGA 759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240

Qy    760 GCTGTCCTGCTGGTGGCCTTTATGAGTAGTGTTCTCTGTCCAGATGTTGAATTCAAACGT 819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArg 260

Qy    820 CGTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAGGGTG 879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlyGlySerTyrMetGluVal 280

Qy    880 GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGAGTGGTGCAT 939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGluValValHis 300

Qy    940 GTGATTGACGTTGACCGGGGAGGAGAGAAGAGGGGAAAGACCAATCTGAGAGGTATTG 999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ValIleAspValAspArgGlyGlyGluLysLysGlyLysAspGlnSerGlyGluValLeu 320

Qy   1000 TCTTCAGTA 1008
          |||||||||
Db    321 SerSerVal 323
```

```
RESULT 2
ADN38936
ID   ADN38936 standard; protein; 323 AA.
XX
AC   ADN38936;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:254.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularisation syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0333393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
```

```
Qy   340 ACTGTGGCCATGGTGTGCACAGGAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA 399
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 ThrValAlaMetValCysThrArgIleSerIleAlaLysSerValPheTyrIleAla 120

Qy   400 GCCCAGTGCCTGGGGCCCATCATTGGACAGGAGGAATCCTATCTGGTCACACCTCCCAGT 459
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer 140

Qy   460 GTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC 519
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160

Qy   520 CTGGTTGAGTTGATAATCACATTCAATTGGTTTACTATCTTTGCCAGCTGTGATTCC 579
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180

Qy   580 AAACGGACTGATGTCACTGGCTCAATAGCTTTAGCAATTGAATTTCTGTTGCAATTGGA 639
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LysArgThrAspValThrGlySerIleAlaLeuAlaLeuAlaIleGluPheLeuLeuGly 200

Qy   640 CATTTATTTGCAATCAATTAATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCT 699
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220

Qy   700 GCAGTTATCATGGGAAATTGGAAAACCATTGGAATATATTGGGTGGGCCCATCATAGGA 759
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240

Qy   760 GCTGTCCTGCTGGTGGCCTTATGAGTATGTCTCTGTCCAGATGTTGAATTCAAACGT 819
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArg 260

Qy   820 CGTTTAAAGAAGCCTTCAGCAAAGTGCCCAGCAACAAAGGAAGCTACATGGAGGTG 879
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   261 ArgPheLysGluAlaPheSerLysAlaAlaAlaGlnThrLysGlySerTyrMetGluVal 280

Qy   880 GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCAT 939
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300

Qy   940 GTGATTGACGTTGACCGGGGAGGAGGAAGAAGAGGGGAAAAGACCAATCTGAGAGGTATTG 999
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ValIleAspValAspArgGlyGluGlyLysLysAspGlnSerGlyGluValLeu 320

Qy  1000 TCTTCAGTA 1008
         |||||||||
Db   321 SerSerVal 323
```

RESULT 3
ADS13763
ID   ADS13763   standard; protein; 323 AA.
XX
AC   ADS13763;
XX
DT   02-DEC-2004 (first entry)
XX
DE   Human aquaporin, AQP4.
XX
KW   Aquaporin; water channel; hypotensive; cardiovascular; hepatotropic;
KW   nephrotropic; endocrine; cytostatic; antiallergic; antibacterial;
KW   immunosuppressive; virucide; fungicide; human; AQP4.
XX
OS   Homo sapiens.
XX
PN   WO2004077010-A2.
XX
PD   10-SEP-2004.
XX
PF   03-FEB-2004; 2004WO-US003022.
XX
PR   03-FEB-2003; 2003US-0444846P.
XX

PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI   Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN38935.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO 254; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and nucleic acids are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularisation syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 323 AA;

Alignment Scores:
```
Pred. No.:
Score:                 2.09e-158     Length:        323
Percent Similarity:    1685.00       Matches:       323
Best Local Similarity: 100.00%       Conservative:  0
Query Match:           100.00%       Mismatches:    0
DB:                    77.26%        Indels:        0
                       7             Gaps:          0
```

US-10-723-180-1 (1-1152) x ADN38936 (1-323)

```
Qy    40 ATGAGTGACGACCCACAGCCAAGCCAAGGCGGTGGGTAAGTGTGACCTTTGTCGTACCAGAGAG 99
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetSerAspArgProThrAlaArgArgTrpGlyLysCysGlyProLeuCysArgThrArgGlu 20

Qy   100 AACATCATGTGGCTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCG 159
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40

Qy   160 GAATTTCTGACCATGCTTATTTTTGTTCTCCTCAGCCTGGATCCACCACTCAACTGGGGT 219
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuAspProProLeuAsnTrpGly 60

Qy   220 GGAACAGAAAAGCCTTACCGGTCCACATGGTTCCATCCCCTTGCTTTGGACTCAGC 279
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80

Qy   280 ATTGCAACCATGGTGCAGTGCTTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG 339
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 IleAlaThrMetValGlnCysPheTrpProTyrGlnArgTrpAlaHisIleAsnProAlaVal 100
```

```
PA   (SCRI ) SCRIPPS RES INST.
PA   (DANI/) DANIELS M.
PA   (YEAG/) YEAGER M.
XX
PI   Daniels M, Yeager M;
XX
DR   WPI; 2004-662063/64.
DR   N-PSDB; ADS13764.
DR   GENBANK; AAH22286.
XX
PT   Identifying agents that can modulate water channel activity, useful for
PT   treating hypertension or congestive heart failure and in affecting the
PT   hydraulic conductivity of other plant tissues, comprises obtaining an
PT   aquaporin-expressing cell.
XX
PS   Claim 3; SEQ ID NO 7; 69pp; English.
XX
CC   The invention relates to identifying agents that can modulate water
CC   channel activity. The method involves obtaining an aquaporin-expressing
CC   cell that is osmotically sensitive in an osmotically acceptable solution;
CC   contacting the aquaporin-expressing cell that is osmotically sensitive
CC   with a test agent; osmotically shocking the aquaporin-expressing cell
CC   that is osmotically sensitive in the presence of the test agent; and
CC   observing whether water channel activity of the aquaporin is modulated.
CC   The identified agents are useful for treating hypertension, congestive
CC   heart failure and clinical states relating to abnormal water balance,
CC   e.g. hepatic cirrhosis, nephrotic syndrome, renal failure, endocrine
CC   disorders, cancers, allergic reactions, or sepsis from viral, bacterial,
CC   or fungal organism, in altering water uptake by plant roots, and in
CC   affecting the hydraulic conductivity of other plant tissues. The present
CC   sequence represents a human aquaporin, AQP4 polypeptide.
XX
SQ   Sequence 323 AA;

Alignment Scores:
Pred. NO.:
Score:                  1685.00
Percent Similarity:     100.00%        Length:        323
Best Local Similarity:  100.00%        Matches:       323
Query Match:             77.26%        Conservative:    0
DB:                          8         Mismatches:      0
                                       Indels:          0
                                       Gaps:            0

US-10-723-180-1 (1-1152) x ADS13763 (1-323)

Qy   40 ATGAGTGACAGACCCACAGCCAAGCGGTGGGTAAGTGTGGACCTTTGTGTACCAGAGAG 99
                                                                        
Db    1 MetSerAspArgProThrAlaAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20

Qy  100 AACATCATGGCCTTCAAAGGGTCTCGGACTCAAGCTTTCTCGGAAAGCAGTCACAGCG 159
                                                                        
Db   21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheThrGlyLysAlaValThrAla 40

Qy  160 GAATTCTGGCCATGCTTATTTTGTTCTCCCTCAGCCTGGGATCCACCATCAACTGGGGT 219
                                                                        
Db   41 GluPheLeuAlaMetLeuLeuPheValLeuLeuSerLeuSerGlySerThrIleAsnTrpGly 60

Qy  220 GGAACAGAAAAGCCTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGC 279
                                                                        
Db   61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80

Qy  280 ATTGCAACCATGGTGCAGTGCTTTGGCCCATATCAGCGGTGGCCACATCACCCTGCAGTG 339
                                                                        
Db   81 IleAlaThrMetValGlnCysPheGlyProTyrHisGlyHisIleAsnProAlaVal 100

Qy  340 ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATGGCCAAGTCTGTCTTCTACATGCCA 399
                                                                        
Db  101 ThrValAlaMetValCysThrArgLysIleSerMetAlaLysSerValPheTyrIleAla 120

Qy  400 GCCCAGTGCCTGGGGCCATCATTGGAGCAGGAATCCTTATCCTGGTCACACCTCCCAGT 459
                                                                        
Db  121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer 140

Qy  460 GTGGTGGGAGGACGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC 519
```

```
Db  141 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160

Qy  520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTGCCAGCTGTGATTCC 579
                                                                        
Db  161 LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180

Qy  580 AAACGGACTGATGTCACTGCTCCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA 639
                                                                        
Db  181 LysArgThrAspValThrGlySerIleAlaLeuAlaLeuIleGlyPheSerValAlaIleGly 200

Qy  640 CATTTATTTGCAATCAATTACTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCT 699
                                                                        
Db  201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyLeuPro 220

Qy  700 GCAGTTATCATGGGAAAATGGGAAAACCATTGGATATATTGGGGTTGGGCCCATCATAGGA 759
                                                                        
Db  221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240

Qy  760 GCTGTCCTCGCTGGTGGCCCTTATGAGTATGTCTTCTGTCCAGATGTGAATTCAAACGT 819
                                                                        
Db  241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArg 260

Qy  820 CGTTTTAAAGAAGCTTCAGCAAAGCTGCCAGCAAACAAAAGGAAGCTACATGGAGGTG 879
                                                                        
Db  261 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 280

Qy  880 GAGGACAACAGCAGGAGTCAGGTAGACGACGATGACCTGATTCTAAAACCTGAGTGGTGCAT 939
                                                                        
Db  281 GluAspAsnArgSerGlnValGluThrAspAspAspLeuIleLeuLysProGlyValValHis 300

Qy  940 GTGATTGACGTTGACCGGGGAGGAGGAGAAGAGGGGAAAGACCAATCTGAGAGGTATTG 999
                                                                        
Db  301 ValIleAspValAspArgGlyGluGluGluLysGlyLysAspGlnSerGlyGluValLeu 320

Qy 1000 TCTTCAGTA 1008
                 
Db  321 SerSerVal 323

RESULT 4
ABB57089
ID   ABB57089 standard; protein; 300 AA.
XX
AC   ABB57089;
XX
DT   07-MAR-2002 (first entry)
XX
DE   Mouse ischaemic condition related protein sequence SEQ ID NO:194.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP004192.
XX
PR   18-MAY-2000; 2000JP-00145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   N-PSDB; ABI99325.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or by
PT   determining the expression profile of a gene group comprising these
PT   genes.
```

```
XX
PS   Claim 2; Page 557-559; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression profile of particular
CC   genes (I) in a test sample or determining the expression profile of a
CC   gene group in the sample comprising genes selected from (I). The method
CC   is useful for examining the ischaemic condition (e.g. compressive
CC   ischaemia, occlusive ischaemia or vasopastic ischaemia) by measuring
CC   expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC   protein sequences in ABB57020 to ABB57374) or by determining the
CC   expression profile of a gene group comprising these genes. The expression
CC   levels or expression profiles produced by these genes are used as an
CC   indicator when screening for ischaemic condition-improving drugs or
CC   therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC   primers for a mouse ischaemic condition related sequence, which are used
CC   in the exemplification of the present invention
XX
SQ   Sequence 300 AA;

Alignment Scores:
Pred. No.:
Score:               7.24e-136       Length:          300
Percent Similarity:  1458.50         Matches:         281
Best Local Similarity: 96.01%        Conservative:    8
Query Match:         93.36%          Mismatches:      11
DB:                  66.87%          Indels:          1
                     5               Gaps:            1

US-10-723-180-1 (1-1152) x ABB57089 (1-300)

Qy    106 ATGGTGGCTTCAAAAGGGGTCTGGACTCAAGCTTTCTGAAAAGCAGTCACAGCGGAATTT 165
          ||||||||||||||||||||||||||||||||||||  |||||||  ||:||||||
.Db     1 MetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValSerAlaGluPhe  20

Qy    166 CTGGCCATGCTATTTTGTTCCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACA 225
          |||||||||||||||  |||  :::||||||||||||||||||||||||||||:::
Db     21 LeuAlaThrLeuIlePheVal---LeuGlyValGlySerThrIleAsnTrpGlyGlySer  39

Qy    226 GAAAAGCCTTACCGGTCGACATGGTTCTCATCTCCTTTGCTTTGGACTCAGCATTGCA 285
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 GluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla  59

Qy    286 ACCATGGTGCAGTGCTTTGGCCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTG 345
          |||||||  ||||||||| ||||||||||||||||||||||||||||||||||||||
Db     60 ThrMetValGlnCysLeuGlyHisIleSerGlyHisGlyHisGlyIleAsnProAlaValThrVal  79

Qy    346 GCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAG 405
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 AlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleAlaAlaGln  99

Qy    406 TGCCTGGGGGCCATCATTGAGCAGGAAATCCTCTATCTGTCCACACCTCCCAGTGTGGTG 465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 CysLeuGlyAlaIleIleGluGlnGluIleLeuTyrLeuSerThrProProSerValVal 119

Qy    466 GGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCCGCTGGTCATGGTCTCCTGGTT 525
          ||||||||||||  |||||||  |||||||||||||||||||||||||||||||||||
Db    120 GlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrArgLeuSerLeuLeuVal 139

Qy    526 GAGTTGATAATCACATTTCAATTGGTGTTACTACTTTTGCCAGCTGTGATTCCAAACCGG 585
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.-Db   140 GluLeuIleIleThrPheGlnLeuValPheThrValPheAlaSerCysAspSerLysArg 159

Qy    586 ACTCATGTCACTGGCTCAATAGCTTTAGCAATTGGGATTTTCTGTTGCAATTGGACATTTA 645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 ThrHisValThrGlySerIleAlaLeuAlaLeuAlaIleGlyGlyPheSerValAlaIleGlyHisLeu 179

Qy    646 TTTGCAATCAATTATACTGGTGCCAGCATCGAATCCCGCCCGATCCTTGGACCTGCAGTT 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaVal 199

Qy    706 ATCATGGGAAATTGGAAAACCATTGAATATTATTGGGTTGGCCCATCATAGAGAGCTGTC 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 IleMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGlyAlaVal 219

Qy    766 CTCGGCTGGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTT 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 LeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArgArgLeu 239

Qy    826 AAAGAAGCCTTCAGCAAGCTGCCCAGCAAACAAAAGAAGCTACATGGAGGTGGGAGGAC 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 LysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAsp 259

Qy    886 AACAGGAGTCAGGTAGAGACGGATGACCTGATTCAAAACCTGGAGTGGTGCCATGTCGATT 945
          |||||||||||||||||||||||||::::||||||||||||||||||||||||||||
Db    260 AsnArgSerGlnValGlnThrGluAspLeuIleLeuLeuLysProGlyValValHisValIle 279

Qy    946 GACGTTGACCCGGGAGGAGGAGGAAGAAGGGGAAAGACCAATCTGGAGAGGTATTGTCTTCA 1005
          |||:|||||||||||||||||||||||||||||||::::|||||||||
Db    280 AspIleAspArgGlyGluGluLysGlyLysArgAspSerSerGlyGluValLeuSerSer 299

Qy   1006 GTA 1008
          ||
Db    300 Val 300
```

```
RESULT 5
ABG96280
ID   ABG96280 standard; protein; 265 AA.
XX
AC   ABG96280;
XX
DT   11-DEC-2002  (first entry)
XX
DE   Human ovarian cancer marker OV3.
XX
KW   Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW   central nervous system disorder; bacterial meningitis; viral meningitis;
KW   Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW   brain herniation; inflammation; encephalitis; testicular disorder;
KW   nontuberculous granulomatous orchitis; connective tissue disorder;
KW   heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW   histological type; carcinogenic; ovarian cancer marker.
XX
OS   Homo sapiens.
XX
PN   WO20071928-A2.
XX
PD   19-SEP-2002.
XX
PF   14-MAR-2002; 2002WO-US007826.
XX
PR   14-MAR-2001; 2001US-0276025P.
PR   14-MAR-2001; 2001US-0276026P.
PR   10-AUG-2001; 2001US-0311732P.
PR   19-SEP-2001; 2001US-0323580P.
PR   26-SEP-2001; 2001US-0324967P.
PR   26-SEP-2001; 2001US-0325102P.
PR   26-SEP-2001; 2001US-0325149P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI   Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI   Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR   WPI; 2002-723277/78.
DR   N-PSDB; ABS76373.
XX
PT   Assessing whether a patient is afflicted with ovarian cancer, useful in
PT   assessing the stage or progression of the disease, comprises comparing
PT   the expression level of a cancer marker in a sample from a patient and
PT   from a non cancer patient.
XX
PS   Disclosure; Page 124-125; 481pp; English.
XX
CC   The present invention relates to a new method for assessing whether a
CC   patient is afflicted with ovarian cancer. The method involves comparing
```

CC   the expression level of a marker in a patient sample and the normal level
CC   of expression of the marker in a control non-ovarian cancer sample, where
CC   the marker is selected from 363 cancer markers described in the
CC   specification. The method of the invention is useful in diagnosing or
CC   characterising cancer, in detecting the presence of cancer as early as
CC   possible, and the recurrence of ovarian cancer. The method may also be of
CC   particular use with patients having an enhanced risk of developing
CC   ovarian cancer (e.g. patients having a familial history of ovarian
CC   cancer). The cancer markers may be used in the management and treatment
CC   of e.g. brain and central nervous system disorders (e.g. bacterial and
CC   viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC   disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC   inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC   testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC   connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC   disease or atherosclerosis). The compositions and methods may also be
CC   used in assessing the histological type of neoplasm associated with
CC   ovarian cancer, monitoring the progression of ovarian cancer, determining
CC   whether ovarian cancer has metastasized or is likely to metastasize,
CC   selecting a composition for inhibiting ovarian cancer, assessing the
CC   ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC   cancer or at risk of developing ovarian cancer. The present amino acid
CC   sequence represents one of the ovarian cancer markers described in the
CC   invention
XX
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:   4.27e-52   Length: 265
Score: 617.00   Matches: 132
Percent Similarity: 64.64%   Conservative: 38
Best Local Similarity: 50.19%   Mismatches: 81
Query Match: 28.29%   Indels: 12
DB: Gaps: 4

.US-10-723-180-1 (1-1152) x ABG96280 (1-265)

Qy   118 AAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTT 177
             ||| ||||| :: |||| ||||||| ||||||||:|||||||||||||
Db   3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheAlaThrLeu   22

Qy   178 ATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGAACAGAAAAGCCTTTA 237
             ||||||||||||||||||| :: |||
Db   23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-------ProSerAlaLeu   39

Qy   238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCAATTGCAACCATGTGCAG 297
             |||| ::::::: |||||:::|||| |||||||||::||||| |||||
Db   40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy   298 TGCTTTGGCCATATCAGCCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
             ||||||||||||||||||||:||||||||||||:||:||:|||||||||:::
Db   59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal 78

Qy   358 ACCAGGAAGATCAGGATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCC 417
             ::||||||| :::::: |||||| | |||||||||||||||||:||||||||||||||
Db   79 GlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAlaGlnLeuVaIGlyAla 98

Qy   418 ATCATTGGAGCAGGAATCCTCTATCGGGTCACACCTCCCAGTGTGGTGGGAGGGCCTGGGA 477
             || ||||||||||||||| ||| :: || ||| ||
Db   99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy   478 GTCACCATGGTTCATGGAAATCTTACCGGCTGGTCATCGTTCTCCGGTTGAGTTGATAATC 537
             ||| ||||| || || |||||:||||||||||||||||||||
Db   119 ValAsnAlaLeuAsnAsnThrThrGlnIleGlyLnAlaMetValValGluLeuIleLeu 138

Qy   538 ACATTTCAATTGGTTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCACT 597
             |||||||||||:||:|| ||||||||:::||||||||||
Db   139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAlaSerProAlaAlaSerProVal 158

Qy   598 GGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTTATTTGCAATCAAT 657
             ||||||||| ||||||:::|||||||||
Db   159 GlySerProAlaLeuSerIleGlyLeuValThrLeuGlyHisLeuValGlyIleLeuTyr 178

Qy   658 TATACTGGTGCCAGCATGAATCCGCCCGATCCTTGGACCTGCCAGTTATCATGGGAAAT 717
             ::|||||| |||||| ||||| ||||||||||||||||||||:::||:|||
Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg   198

Qy   718 TGGGAA--AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCTGGT 774
             || :   ::||||||||| ||||:::||||||||||||:::||||||
Db   199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla 218

Qy   775 GGCCTTTATGAGTATGTCTTCTGTCCAGATGTGAATTCAAACCTCGTTTTAAGAGAAGCC 834
             || ::: ||||:::|||||:|| |::
Db   219 IleLeuTyrPheTyrLeuLeuPheProAsn-----------------SerLeuSer 231

Qy   835 TTCAGCAAGCTGCCCAGCAAACAAAAGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
             |||:: ||| ::|||||||||||||||||||| |:|
Db   232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspTrpGluGluGlu 251

Qy   895 CAGGTAGAG 903
             ||| |||
Db   252 GlnArgGlu 254

RESULT 6
ABB97500
ID   ABB97500 standard; protein; 265 AA.
XX
AC   ABB97500;
XX
DT   27-JUN-2002 (first entry)
XX
DE   Novel human protein SEQ ID NO: 768.
XX
KW   Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW   antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW   neuroprotective; antiparkinsonian; protein therapy; EST;
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO20022660-A2.
XX
PD   21-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US026015.
XX
PR   11-SEP-2000; 2000US-00659671.
XX
PA   (HYSB-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI   Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR   WPI; 2002-292408/33.
DR   N-PSDB; ABN32686.
XX
PT   An isolated polynucleotide for treating diseases associated with its
PT   encoded polypeptide such as cancer and multiple sclerosis.
XX
PS   Claim 20; SEQ ID NO 768; 509pp; English.
XX
CC   The present invention provides the protein and coding sequences of 444
CC   novel human proteins. These were isolated from expressed sequences tags
CC   (ESTs). They can be used to stimulate cell growth, to regulate
CC   haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC   e.g. in burn treatment, to regulate the immune system e.g. to treat
CC   multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC   infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC   and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC   rheumatoid arthritis, and to treat nervous system disorders e.g.
CC   Parkinson's disease. The present sequence is a protein of the invention
XX
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:   4.27e-52   Length: 265

```
Score:                617.00    Matches:         132
Percent Similarity:   64.64%    Conservative:     38
Best Local Similarity: 50.19%   Mismatches:       81
Query Match:          28.29%    Indels:           12
DB:                             Gaps:              4

US-10-723-180-1 (1-1152) x ABB97500 (1-265)

Qy    118 AAAGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACACGCGGAATTTCTGGCCATGCTT 177
              ||| ||||||      |||||||||      |||||||||||||||||
Db      3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheAlaThrLeu  22

Qy    178 ATTTTTGTTCCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTA 237
              ||||||||      ||||||||       ::: |||
Db     23 IlePheValPheGlyLeuGlySerAlaLeuLysTrp--------ProSerAlaLeu  39

Qy    238 CCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAG 297
              |||      |||||||      ::::: |||||||| ||||||||||||||||
Db     40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln  58

Qy    298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
              ||::||||||||      ||| ::|||||||||||||| ||:::|||||||:::::::
Db     59 AlaLeuGlyProValSerGlyValGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal  78

Qy    358 ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCACATCGCAGCCCCAGTGCCTGGGGGCC 417
              ::||||||||::::   :::::: ||||||||      ::||||||    ::||||||||
Db     79 GlyAsnGlnIleSerLeuLeuArgAlaPheThrValAlaAlaAlaGlnLeuLeuValGlyAla  98

Qy    418 ATCATTGGAGCAGGAATCCTATCTGGTCACACCTCCCAGTGTGGTGGAGGCCTGGGA 477
              ||||||||||::||||      :::::: ||||      ::||||| ||||||||
Db     99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla  118

Qy.   478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTCCTGGTTGAGTTGATAATC 537
              ||      |||||||||      ||| |||      ||||||::::||||||||||||||
Db    119 ValAsnAlaLeuAsnAsnThrThrGlnIleGlyAsnAlaMetValGlnLeuIleLeuLeu  138

Qy    538 ACAATTCAATTGGTGTTTACTATCTTGCCAGCTGTGATTCCAAACGGACTGATGTCACT 597
              ||||||||||      ||||||      |||||||::::|||||
Db    139 ThrPheGlnLeuAlaLeuCysIlePheLeuAlaSerThrAspSerArgArgThrSerProVal  158

Qy    598 GGCTCAATAGCTTAGCAATTGGATTTTCTGTTGACAATTTATTTGCAATCAAT 657
              |||||||      ||| ||||||||||| ||||||      ||||||
Db    159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr  178

Qy    658 TATACTGGTGCCAGCATGAATCCGCCGGCCGATCCTTTGACCTCCAGTTATCATGGGAAAT 717
              ::||||||      ||||      ||||      ||::::||||||||||||||||||:||
Db    179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg  198

Qy    718 TGGGAA---AACCATTGGATATATTGGGGTTGGGCCCATCATAGAGAGCTGTCCTCCGTGGT 774
              ::: ||||||||||      |||||::::|||||||      ||||||||
Db    199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyValAlaValLeuAlaAla  218

Qy    775 GGCCTTTATGAGTATGTCTTCGTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
              |||||||      |||||||||||||      |::::|||
Db    219 IleLeuTyrPheTyrLeuLeuPheProAsn-----------SerLeuSer  231

Qy    835 TTCAGCAAGCTGCCCAGCAAACAAAGAAGCTACATGGAGGTGGAGGACAACAGCAGAGT 894
              ||||||      :::::||||||::||| |||||||
Db    232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspGluAspTrpGluGlu  251

Qy    895 CAGGTAGAG 903
              |||||:|
Db    252 GlnArgGlu 254
```

RESULT 7
ABR92054

```
ID   ABR92054 standard; protein; 265 AA.
XX
AC   ABR92054;
XX
DT   10-SEP-2003  (first entry)
XX
DE   Human cervical cancer cell marker protein SEQ ID NO:16.
```

Alignment Scores:

```
Pred. No.:                       4.27e-52    Length:          265
Score:                           617.00      Matches:         132
Percent Similarity:              64.64%      Conservative:     38
Best Local Similarity:           50.19%      Mismatches:       81
Query Match:                     28.29%      Indels:           12
DB:                                          Gaps:              6

US-10-723-180-1 (1-1152) x ABR92054 (1-265)

Qy    118 AAAGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACACGCGGAATTTCTGGCCATGCTT 177
              ||| ||||||      |||||||||      |||||||||||||||||
Db      3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheAlaThrLeu  22

Qy    178 ATTTTTGTTCCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTA 237
              ||||||||      ||||||||       ::: |||
Db     23 IlePheValPheGlyLeuGlySerAlaLeuLysTrp--------ProSerAlaLeu  39

Qy    238 CCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAG 297
              |||      |||||||      ::::: |||||||| ||||||||||||||||
Db     40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln  58

Qy    298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
              ||::||||||||      ||| ::|||||||||||||| ||:::|||||||:::::::
Db     59 AlaLeuGlyProValSerGlyValGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal  78
```

```
XX
DR   WPI, 2003-268312/26.
DR   GENBANK; NP_001642.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Example 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (described in Table 3
CC   of the specification) which is differentially expressed during pain.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:
Score:                    4.27e-52      Length:        265
Percent Similarity:       617.00        Matches:       132
Best Local Similarity:    64.64%        Conservative:   38
Query Match:              50.19%        Mismatches:     81
DB:                       28.29%        Indels:         12
                          7             Gaps:            4

US-10-723-180-1 (1-1152) x ADD47950 (1-265)

Qy  118 AAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCAGCGGAATTTCTGGCCATGCTT 177
            ||| |||  :::   ||||| |||||||  ||||||| |||||||:||
Db    3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheLeuAlaThrLeu 22

Qy  178 ATTTTTGTTCTCCTGGATCCACATCAACTGGGTGGAACAGAAAAGCCTTTA 237
            ||:|||||:::||||||||||:::||||||||||||-------ProSerAlaLeu 39
Db   23 IlePheValPheGlyLeuGlySerAlaLeuLysTrp-------ProSerAlaLeu 39

Qy  238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAACCATGTCAACCATGTGCAG 297
            |||    :::||||||||:::||   ||||||||||||| |||
Db   40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy  298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGACTGTGCCATGGTGTGC 357
            |||||||||||| |||||| |||||| :::||||||||||||||||::::||
Db   59 AlaLeuGlyProValSerGlyValGlyHisIleAsnProAlaIleThrLeuLeuLeuVal 78

Qy  358 ACCAGGAGGATCAGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCC 417
            :::|||||||:::       :::||||| ||||||||  :::||||||  |||
Db   79 GlyAsnGlnIleSerLeuLeuArgAlaPheTyrValAlaGlnAlaGlyAsnValGlyAla 98

Qy  418 ATCATTGGAGCAGGAAATCCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTGGGA 477
            |||||||||||||| |||  ||:::||||||||||||  :::|||||
Db   99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy  478 GTCACCATGTTCATGGAAATCTTACCGCTGGTCATGGTCTCCGGTGTGAGTTGATAATC 537
            ||                :::|||         |||||:::||||||
Db  119 ValAsnAlaLeuAsnAlaAsnThrThrGlnIleGlyAlaMetValValGluLeuLeuLeu 138

Qy  538 ACATTTCAATTGGTGTTTACTATCTTGCCAGCTGTGATTCCAAACGGACTGATGTCACT 597
            |||||||         :::          :::||||||||||
Db  139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgArgThrSerProVal 158

Qy  598 GGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTTGCAATCAAT 657
            |||||||       |||:::||||||        ::||||||:::
Db  159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy  658 TATATCTGGTGCCAGGATGGAATCCCGCCGGATCCTTGGACCTGCAGTTATCATGGGAAAT 717
            ::::|||      |||||||||||||||||:::||
Db  179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198

Qy  718 TGGGAA---AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGAGTGTCCTCGCTGGT 774
            :::          |||||:::|||||||||||||||:::|||||
Db  199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla 218

Qy  775 GGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGCC 834
            ||||||||         |||||||||         :::
Db  219 IleLeuTyrPheTyrLeuLeuPheProAsn----------SerLeuSer 231

Qy  835 TTCAGACAAAGCTGCCCAGCAAACAAAGGAGCTACATCGGAGGTGGAGGACACAGGAGT 894
            ||||:::|||                       |||:::|||
Db  232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspTrpGluGlu 251

Qy  895 CAGTAGAG 903
            ||| |||
Db  252 GlnArgGlu 254


RESULT 8
ADD47950
ID   ADD47950 standard; protein; 265 AA.
XX
AC   ADD47950;
XX
DT   02-DEC-2004 (revised)
DT   29-JAN-2004 (first entry)
XX
DE   Human Protein NP_001642, SEQ ID NO 13646.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
OS   Unidentified.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO) GEN HOSPITAL CORP.
PA   (FARB) BAYER AG.
XX
PI   Woolf C, D'urso D, Befort K, Costigan M;
```

CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
CC   injury (CCI) and spared nerve injury (Chung), chronic constriction
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:   Length: 265
Score: 4.27e-52   Matches: 132
Percent Similarity: 617.00   Conservative: 38
Best Local Similarity: 64.64%   Mismatches: 81
Query Match: 50.19%   Indels: 12
DB: 28.29%   Gaps: 4
                             7

US-10-723-180-1 (1-1152) x ADE62232 (1-265)

Qy 118 AAAGGGGTCTGGACTCAAGCTTCTCGGAAAGCAGTCACAGCGGAATTTCTCGGCCATGCTT 177
            ||| ::: ||||||||||| |||||||
Db   3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaLeuAlaThrLeu 22

Qy 178 ATTTTTGTTCCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAAGCCTTTA 237
            |||||||| ||| ::: |||
Db   23 IlePheValPheGlyLeuGlySerAlaLeuLysTrp------ProSerAlaLeu 39

Qy 238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTGCAACCATGGTTGCAG 297
            |||||||||||||| ::: |||||||||||||||
Db   40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy 298 TGCTTTGGCCATATCAGCGGTGGCCACCATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
            ||||:::||||||||||||||||| :::|||||
Db   59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal 78

Qy 358 ACCAGGAAGATCAGCATGCGCCAAGTCGTCTTCTACATCGGCAGCCCAGTGCTGCGGGGGCC 417
            :::|||||| ::::::||||||| |||:::|||||
Db   79 GlyAsnGlnIleSerLeuLeuArgGlnAlaPheGlyTyrValAlaAlaGlnLeuValGlyAla 98

Qy 418 ATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCCAGTGTGTGGAGGCCCTGGGA 477
            ||||||||||||| |||||:::   :::
Db   99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy 478 GTCACCATGGTTCATGGAAATTTCATGGTCTTCACCGCTGGTCATGGTCCTGGTTGAGTTGATAATC 537
            ||| ::::: |||||::::: ||||:::
Db 119 ValAsnAlaLeuAsnAsnThrThrGlnGlnAlaMetValValGluLeuIleLeu 138

Qy 538 ACATTTCAATTGGTGTTCTACTATCTTGCCAGTCTGTGATTCCAAACGGACTGATGTCACT 597
            |||||||||| |||| :::|||||||
Db 139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgThrSerProVal 158

Qy 598 GGCTCAATAGCTTAGCAATTGGATTTTCGTTGCAATGGACATTTATTTGCAATCAAT 657
            |||||||||:::   :::|||||||||
Db 159 GlySerProAlaLeuSerIleLeuGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy 658 TATACTGGTGGTCCAGCATGAATCCCGCCCCGATCCTTGGACCTGCAGTTATCATGGGGAAAT 717
            :::|| |||||||||||||||:::|||
Db 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198

Qy 718 TGGGAA--AACCATTGGATAATATTGGGTGGGCCCATCATAGGAGCTGTCCTCGCTGGT 774
            ::: ||||||:::|||AAGGAGCTGTC
Db 199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla 218

Qy 775 GGCCTTTATGATGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGAGCC 834
            |||||||:::   :::
Db 219 IleLeuTyrPheTyrLeuLeuPheProAsn-----------SerLeuSer 231

Qy 835 TTCAGCAAAGCTGCCCAGCAAACAAAAGAAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
            ||||||| |||||:::|||
Db 232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspGluTrpGluGlu 251

Qy 895 CAGGTAGAG 903
            |||:::
Db 252 GlnArgGlu 254

RESULT 9
ADE62232
ID   ADE62232 standard; protein; 265 AA.
XX
AC   ADE62232;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Protein NP_001642, SEQ ID NO 8161.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO200301 6475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO) GEN HOSPITAL CORP.
PA   (FARB) BAYER AG.
XX
PI   Woolf C, D'urso D, Befort K, Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; NP_001642.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,

```
CC    an animal model for bronchial asthma or chronic obstructive pulmonary
CC    disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC    method for producing an animal model for bronchial asthma or chronic
CC    obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC    asthma or chronic obstructive pulmonary disease, comprising the compound,
CC    a marker gene or an antisense nucleic acid corresponding to a portion of
CC    the marker gene, a ribozyme, a polynucleotide that suppresses the
CC    expression of the gene through an RNAi effect or an antibody recognising
CC    a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC    bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC    probe has been immobilised to assay a marker gene. (1) has respiratory
CC    and antiasthmatic activities, and can be used in gene therapy. The method
CC    is useful for testing for or screening for a therapeutic agent for
CC    bronchial asthma or chronic obstructive pulmonary disease. The present
CC    sequence is used in the exemplification of the present invention.
XX
SQ    Sequence 265 AA;

Alignment Scores:
Pred. No.:                    Length:              265
Score:             4.27e-52   Matches:             132
Percent Similarity: 617.00    Conservative:         38
Best Local Similarity: 64.64% Mismatches:           81
Query Match:       50.19%     Indels:               12
                   28.29%     Gaps:                  4
DB:                8

US-10-723-180-1 (1-1152) x ADJ75631 (1-265)

Qy   118 AAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCAGACGCGAATTTCTGGCCATGCTT 177
             ||||||     :::    ||||| |||||| |||||||||||||||||||||||||
Db     3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheLeuAlaThrLeu 22

Qy   178 ATTTTTGTCTCCTGGCCTGGATCCACCATCAACTGGGGTGAACAGAAAAGCCTTTA 237
             |||||      ::  ||| ||||            :: |||  -------ProSerAlaLeu 39
Db    23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-------ProSerAlaLeu 39

Qy   238 CCGGTCGACATGGTTCTCATCTCCCTTGCTTTGGACTTGCAGCCATTGCAACCATGGTGCAG 297
             ||||||  :::::|||||||:::||||||||||||||||||||
Db    40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy   298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
             ::||||||||||||      ||::||::||||||||||||||
Db    59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal 78

Qy   358 ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCACATCGCAGCCCAGTGCCTGGGGGCC 417
             ::|||||||||||     |||:||||:||||||||||||||||||
Db    79 GlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAlaGlnLeuuValGlyAla 98

Qy   418 ATCATTGGAGCAGGAATCCTCTATCTGGTGCACACCTCCCAGTGTGGTGGGAGGCCTGGGA 477
             ||||||||||  :::       ||| ||||||||||||||||
Db    99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy   478 GTCACCATGGTTCATGAGAAATCTACCGCTGGATGTCTCCTGGTTGAGTTGATAATC 537
             |||       ::  ||| ||||| ::::|||||||||||||||||
Db   119 ValAsnAlaLeuAsnAsnAsnThrThrGlnGlyGlnAlaMetValValGluLeuLeu 138

Qy   538 ACATTTCAATTGGTTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCACT 597
             :::||||||      |||||      :::||||||||||||:::||
Db   139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgThrSerProVal 158

Qy   598 GGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTTGCAATCAAT 657
             |||||||:::       :::||||||:::||:||||||||||||
Db   159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy   658 TATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGGCCAGTTATCATGGGAAAT 717
             :::|||||||||||||       |||||||||||||||||||||
Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198

Qy   718 TGGGAA--AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTGCTGGT 774
             :::       |||||:::|||||||||:||||||||||:::|||||
Db   199 PheSerProAlaHisTrpValPheProValGlyProIleValGlyAlaValLeuAlaAla 218

Qy   775 GGCCCTTTATGAGTATGTCTTCTGCTGCCAGATGTGAATTCAAACGTCGTTTAAAGAAGCC 834
```

```
Qy   658 TATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGGCCAGTTATCATGGGAAAT 717
             :::||||||     ||||||         ||||||||||||||||||
Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198

Qy   718 TGGGAA--AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTGCTGGT 774
             |||:::      |||||:::||||||||:||||||||||||:::|||
Db   199 PheSerProAlaHisTrpValPheProValGlyProIleValGlyAlaValLeuAlaAla 218

Qy   835 TTCAGCAAAGCTGCCCAGCAACAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
             ||||:::       ||||          |||||||
Db   219 IleLeuTyrPheTyrLeuLeuPheProAsn-------------------SerLeuSer 231

Qy   895 CAGGTAGAG 903
             |||:::
Db   232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspGluAspTrpGluGlu 251

Db   252 GlnArgGlu 254

RESULT 10
ADJ75631
ID   ADJ75631 standard; protein; 265 AA.
XX
AC   ADJ75631;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Marker gene related amino acid sequence SEQ ID NO:883.
XX
KW   bronchial asthma; chronic obstructive pulmonary disease;
KW   respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW   gene therapy; marker.
XX
OS   Homo sapiens.
XX
PN   EP1394274-A2.
XX
PD   03-MAR-2004.
XX
PF   04-AUG-2003; 2003BP-00254857.
XX
PR   06-AUG-2002; 2002JP-00229312.
PR   20-MAR-2003; 2003JP-00077212.
XX
PA   (GENO-) GENOX RES INC.
XX
PI   Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX
DR   WPI; 2004-193155/19.
XX
PT   Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT   comparing the expression level of a marker gene in a biological sample
PT   from a subject, comparing the expression level of the gene in a sample from a
PT   healthy subject.
XX
PS   Example 11; SEQ ID NO 883; 241pp; English.
XX
CC   The present invention describes a method of testing for bronchial asthma
CC   or chronic obstructive pulmonary disease. The method comprises
CC   determining the expression level of a marker gene in a biological sample
CC   from a subject, comparing the expression level determined with the
CC   expression level of the marker gene in a biological sample from a healthy
CC   subject, and judging whether the subject has bronchial asthma or chronic
CC   obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC   genes (S1) whose expression levels increase when respiratory epithelial
CC   cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC   whose expression levels decrease when respiratory epithelial cells are
CC   stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC   testing for bronchial asthma or chronic obstructive pulmonary disease;
CC   (2) a kit for screening for a candidate compound for a therapeutic agent
CC   to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
```

```
Db     219  IleLeuTyrPhePheTyrLeuLeuPheProAsn----------SerLeuSer 231
            |||||          |||::                         :::
Qy     835  TTCAGCAAAGTGCCCAGCAAACAAAGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
            |||:::        ||||||::||          |||||
Db     232  LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspTrpGluGlu 251

Qy     895  CAGGTAGAG 903
            |||  ||||
Db     252  GlnArgGlu 254

RESULT 11
ADJ75701
ID   ADJ75701 standard; protein; 265 AA.
XX
AC   ADJ75701;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Marker gene related amino acid sequence SEQ ID NO:953.
XX
KW   bronchial asthma; chronic obstructive pulmonary disease;
KW   respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW   gene therapy; marker.
XX
OS   Homo sapiens.
XX
PN   EP1394274-A2.
XX
PD   03-MAR-2004.
XX
PF   04-AUG-2003; 2003EP-00254857.
XX
PR   06-AUG-2002; 2002JP-00229312.
PR   20-MAR-2003; 2003JP-00077212.
XX
PA   (GENO-) GENOX RES INC.
XX
PI   Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H,  Izuhara K;
XX
DR   WPI; 2004-193155/19.
XX
PT   Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT   comparing the expression level of a marker gene in a biological sample
PT   from a subject, comparing the expression level of the gene in a sample from a
PT   healthy subject.
XX
PS   Example 11; SEQ ID NO 953; 241pp; English.
XX
CC   The present invention describes a method of testing for bronchial asthma
CC   or chronic obstructive pulmonary disease. The method comprises
CC   determining the expression level of a marker gene in a biological sample
CC   from a subject, comparing the expression level determined with the
CC   expression level of the marker gene in a biological sample from a healthy
CC   subject, and judging whether the subject has bronchial asthma or chronic
CC   obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC   genes (S1) whose expression levels increase when respiratory epithelial
CC   cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC   whose expression levels decrease when respiratory epithelial cells are
CC   stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC   testing for bronchial asthma or chronic obstructive pulmonary disease;
CC   (2) a kit for screening for a candidate compound for a therapeutic agent
CC   to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC   an animal model for bronchial asthma or chronic obstructive pulmonary
CC   disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC   method for producing an animal model for bronchial asthma or chronic
CC   obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC   asthma or chronic obstructive pulmonary disease, comprising the compound,
CC   a marker gene or an antisense nucleic acid corresponding to a portion of
CC   the marker gene, a ribozyme, a polynucleotide that suppresses the
CC   expression of the gene through an RNAi effect or an antibody recognising
CC   a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC   bronchial asthma or chronic obstructive pulmonary disease, on which a
CC   probe has been immobilised to assay a marker gene. (I) has respiratory
CC   and antiasthmatic activities, and can be used in gene therapy. The method
CC   is useful for testing for or screening for a therapeutic agent for
CC   bronchial asthma or chronic obstructive pulmonary disease. The present
XX   sequence is used in the exemplification of the present invention.
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:                         Length:              265
Score:                 4.27e-52    Matches:             132
Percent Similarity:    617.00      Conservative:         38
Best Local Similarity: 64.64%      Mismatches:           81
Query Match:           50.19%      Indels:               12
DB:                    28.29%      Gaps:                  4

US-10-723-180-1 (1-1152) x ADJ75701 (1-265)

Qy  118  AAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCCGAATTTCTGGCCATCTT 177
         |||  ||   ::      ||||||| |||||||| |||||||
Db    3  LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheLeuAlaThrLeu 22

Qy  178  ATTTTTGTTCTCCTCAGCCTGGATCCACCATCAACTGGGTGGAACAGAAAAGCCTTTA 237
         ||| :::        || :::                ------
Db   23  IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp------ProSerAlaLeu 39

Qy  238  CCGGTCGACATGGTTCTCATCTCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAG 297
         |||          :::::    |||||||            |||::|
Db   40  ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuIleGlyThrLeuAlaGln 58

Qy  298  TGCTTTGGCCATATCAGGCGGTGGCCACATCAACCCTGCAGTGACTGGTGGCCATGGTGC 357
         |||            :::|||||::||||       |||::||
Db   59  AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal 78

Qy  358  ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGGCAGCCCAGTGCTGGGGGCC 417
         |||       :::       || :::::      ||||||
Db   79  GlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAlaGlnLeuValGlyAla 98

Qy  418  ATCATTGGAGCAGGAATCCTCTATCGGTCACACCTCCCAGTGTGGTGGAGGCCTGGGA 477
         |||||||||||||||||                :: :::
Db   99  IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy  478  GTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTCCTCGGTTGAGTTGATAATC 537
         |||:::         |||                  :::::||||||||||||||
Db  119  ValAsnAlaLeuAsnAsnThrThrGlnGlyGlnAlaMetValValGluLeuIleLeu 138

Qy  538  ACATTTCAATTGGTGTGTTTACTATCTTTGCCAGCTGTGATCTCCAAACGGACTGATGTCACT 597
         |||::|||                      :::||
Db  139  ThrPheGlnLeuValLeuCysIlePheAlaSerThrAspSerArgArgThrSerProVal 158

Qy  598  GGCTCAATAGCTTAGCAATTGATTTTCTGTGCAATTGGACATTATTTGCAATCAAT 657
         |||||||||||||||||::||||||::        |||||||||||||
Db  159  GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy  658  TATACTGGTGCGCAGCATGAAGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAAT 717
         :::||||||         :::||||||||||||||         :::||
Db  179  PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198

Qy  718  TGGGAA---AACCATGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCTGGT 774
         :::                     :::::||||||::||
Db  199  PheSerProAlaHisTrpValPheThrValGlyProIleValGlyValAlaValLeuAlaAla 218

Qy  775  GGGCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
         |||||||||:::|||   |||:::         :::----------
Db  219  IleLeuTyrPheTyrLeuLeuPheProAsn----------SerLeuSer 231

Qy  835  TTCAGCAAAGTGCCCAGCAAACAAAGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
         |||:::        ||||||::||          |||||
Db  232  LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspTrpGluGlu 251

Qy  895  CAGGTAGAG 903
         |||  ||||
Db  252  GlnArgGlu 254
```

```
RESULT 12
ADL70563
ID   ADL70563 standard; protein; 265 AA.
XX
AC   ADL70563;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Cervical cancer marker OV3A, aquaporin 5 variant 1.
XX
KW   Human; cervical cancer; marker; diagnosis; gene therapy; aquaporin 5;
KW   vaccine; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004018999-A2.
XX
PD   04-MAR-2004.
XX
PF   20-AUG-2003; 2003WO-US026184.
XX
PR   20-AUG-2002; 2002US-0404770P.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Monahan JB, Zhao X, Chen Y, Glatt K, Kamatkar S;
XX
DR   WPI; 2004-248103/23.
DR   N-PSDB; ADL70562.
XX
PT   Assessing, characterizing, monitoring, preventing and treating cervical
PT   cancer including pre-malignant conditions, e.g. dysplasia comprises
PT   comparing the level of expression of a marker in a patient sample and in
PT   a control sample.
XX
PS   Disclosure; SEQ ID NO 6; 167pp; English.
XX
CC   The present sequence is the protein sequence of cervical cancer marker
CC   OV3A, characterised as human aquaporin 5 variant 1. A higher than normal
CC   expression level of this marker correlates with the presence of cervical
CC   cancer, including pre-malignant conditions such as dysplasia, in a
CC   patient. The marker was identified by transcription profiling experiments
CC   and by in situ hybridisation. The invention relates to newly discovered
CC   cancer marker ADL70558-ADL70601 polynucleotides and polypeptides
CC   associated with the cancerous state of cervical cells. It provides
CC   methods, reagents and kits for diagnosing, staging, prognosing,
CC   monitoring and treating cervical cancer (including carcinoma in situ,
CC   invasive carcinoma and metastatic carcinoma) and pre-malignant conditions,
CC   (such as dysplasia, including CIN or SIL). A claimed method of treating
CC   cervical cancer uses an antisense oligonucleotide that is complementary
CC   to one of the cancer marker polynucleotides. Claimed methods of
CC   inhibiting cervical cancer comprise administering a composition that
CC   induces a lower expression of a marker or inhibiting expression of a
CC   marker gene.
XX
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:                              Length:        265
Score:              4.27e-52            Matches:       132
                    617.00             Conservative:   38
Percent Similarity:  64.64%            Mismatches:     81
Best Local Similarity: 50.19%          Indels:         12
Query Match:        28.29%             Gaps:            4
DB:                     8
```

```
Db           23 IleFheValPhePheGlyLeuGlySerAlaLeuLysTrp--------ProSerAlaLeu  39
                |||||||||                 |||||||||           :::        |||

Qy          238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTTGCAGCATTGCAACCATGGTGCAG  297

Db           40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln  58
                |||      :::::::|||||||||||||||||||||:::::

Qy          298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGC  357

Db           59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal  78
                |||        :::|||||||||||||||||||||||||||||||||||||||||||||

Qy          358 ACCAGGAAGATCAGCATGCCAAGTCTGTCTTCTACATCGCAGCCAGTGCCTGGGGGCC  417

Db           79 GlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAlaGlnLeuValGlyAla  98
                :::|||||||||                ||||||||||:::::::|||||||||

Qy          418 ATCATTGGAGCAGGAAATCCTCTATCTGGTCACACTCCCAGTGTGGTGGGAGGCCTGGGA  477

Db           99 IleAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla  118
                |||||||||||        :::       |||||||||:::

Qy          478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTCCTGGTTGAGTTGATAATC  537

Db          119 ValAsnAlaLeuAsnAsnAsnThrThrGlnGlyGlnAlaMetValValGlyLeuLeuLeu  138
                |||        :::|||       ||| :::|||||||||:::::::|||||||||||||

Qy          538 ACATTCAATTGGTGTTTACTATCTTTGCCAGCTGTGATTCCAAAACGGACTGATGTGTCACT  597

Db          139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgArgThrSerProVal  158
                :::|||||||||||||           :::|||||||||

Qy          598 GGCTCCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTTGCAATCAAT  657

Db          159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr  178
                |||||:::|||||||||:::       |||||||||||||:::

Qy          658 TATACTGGTGCCAGCATGAATCCCGCCCATCCTTTGGACCTGCAGTTATCATGGGAAAT  717

Db          179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg  198
                ::::|||||       |||||||||||||||||:::::

Qy          718 TGGGAA---AACCATTGGATATATTGGGTWGGGCCCATCATAGGAGCTGTCCTCGCTGGT  774

Db          199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla  218
                |||::::

Qy          775 GGGCCTTTATGAGTATGTCTTCTGTCCAGATGTGAATTCAAACGTCGTTTAAAGAAGCC  834

Db          219 IleLeuTyrPheTyrLeuLeuPheProAsn------------------SerLeuSer  231
                |||||:::       ||||||

Qy          835 TTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGT  894

Db          232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspGluAspTrpGluGlu  251
                |:::::               |||||

Qy          895 CAGTAGAG  903

Db          252 GlnArgGlu  254
```

```
RESULT 13
AAW55787
ID   AAW55787 standard; protein; 265 AA.
XX
AC   AAW55787;
XX
DT   14-JUL-1998  (first entry)
XX
DE   Rat aquaporin-5.
XX
KW   Human; aquaporin-1; AQP-1; water channel protein; regulation;
KW   osmotic change; erythrocyte; dryness; blindness; hydration; asthma;
KW   secretion; aquaporin-5; AQP-5.
XX
OS   Rattus sp.
XX
PN   US5741671-A.
XX
PD   21-APR-1998.
XX
PF   06-JUN-1995;         95US-00468763.
XX
```

```
US-10-723-180-1 (1-1152) x ADL70563 (1-265)

Qy          118 AAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCAGCGGAATTTCTGGCCATGCTT  177
                 |||    |||    |::    ||||||||||||||||||||||||||||||||||     |||

Db            3 LysGluValCysSerValAlaPheLeuArgValCysAlaValPheAlaGluGluAlaThrLeu  22

Qy          178 ATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTCGGGTGGTGAACTGGGTGAACAGAAAAGCCTTTA  237
```

PR   17-AUG-1992;   92US-00930168.
PR   24-FEB-1995;   95US-00393996.
XX
PA   (UYJO) UNIV JOHNS HOPKINS.
XX
PI   Agre PC;
XX
DR   WPI; 1998-260501/23.
DR   N-PSDB; AAV25906.
XX
PT   Polynucleotide(s) encoding water channel protein Aquaporin-1 - are useful
PT   for recombinant production of protein for activity studies.
XX
PS   Example 9; Fig 14; 48pp; English.
XX
CC   The present sequence represents aquaporin-5 (AQP-5), isolated from a rat
CC   submandibular gland cDNA library, which is used in an example of the
CC   present invention. This invention describes AQP-1 isolated from
CC   human erythrocytes. AQP-1 is a water channel protein. Water channel
CC   proteins regulate the passage of water in and out of cells, in response
CC   to osmotic changes. The DNA encoding AQP-1 is useful for the recombinant
CC   production of AQP-1, found in mammalian erythrocytes, and is useful in
CC   the study to identify reagents which enhance or inhibit water channel
CC   function. This can lead to therapeutics which enhance secretion e.g. in
CC   the case of dryness of eyes which can lead to blindness or to hydrate
CC   large respiratory airways, as their dryness can precipitate asthma
XX
SQ   Sequence 265 AA;

Alignment Scores:
Pred. No.:
Score: 8.5e-52   Length: 265
                        614.00   Matches: 133
Percent Similarity: 56.83%   Conservative: 46
Best Local Similarity: 42.22%   Mismatches: 82
Query Match: 28.15%   Indels: 54
DB: 2   Gaps: 5

US-10-723-180-1 (1-1152) x AAW55787 (1-265)

```
Qy   118 AAAGGGGTCTGGACTCAAGCTTTCTCGAAAGCAGTCACAGCGGAATTCTGGCCATGCTT 177
             |||    ::  |||  :::|||||||  ||||||||||||
Db     3 LysGluValCysSerLeuAlaPhePheLysGluAlaValPheAlaGluPheAlaThrLeu 22

Qy   178 ATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTA 237
         ||||||||||||  ::  ||
Db    23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-------ProSerAlaLeu 39

Qy   238 CCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAG 297
         ||    :::::   |||:::    ||||||:::|    |||:::||  ||||:::|
Db    40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy   298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGTGTGC 357
         ||          :::||||   |||:::||||||:::|||:::::::
Db    59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeueuIle 78

Qy   358 ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGCC 417
         ||||||||||||     :::||||     ||:::|||:::::
Db    79 GlyAsnGlnIleSerLeueuArgGluAlaValPheTyrValAlaAlaGlnLeuValGlyAla 98

Qy   418 ATCATTGGAGCAGGAATCCTATCTGGTCACACCTCCAGTGTCGTGGGGAGGCCTGGGA 477
         ||||||||||||        :: |||  :: |||  ||
Db    99 IleAlaGlyAlaGlyIleLeuTyrTrpLeuAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy   478 GTCACCATGGTTCATGGAAATCTTACCGCGTGGTCATGGTCTCCTGGTTGATAATC 537
         ||    :::::   |||  ||| :::||||:::   :::||||||||||:::
Db   119 ValAsnAlaLeuAsnAsnAsnThrThrProGlyLysAlaMetValValGluLeuIleLeu 138

Qy   538 ACATTTCAATTGGTGTTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCACT 597
         ||||||   ||| ||| :::   |||::|||||:::|||    ||| :::||||
Db   139 ThrPheGlnLeuAlaLeuCysIlePheSerSerThrAspSerArgArgThrSerProVal 158

Qy   598 GGCTCAATAGCTTAGCATTGGATTTTCTGTTGCAATGACATTTATTTGCAATCAAT 657
         ||||||     ||||| :::||||||
```

```
Db   159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178
Qy   658 TATACTGGTGCCAGCATGAATCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAAT 717
         :::|||||    |||:||:::|||||||||||:::||:::||
Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198
Qy   718 TGGGA--AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCGTGTCCTCGCTGGT 774
         :::     :::||||||:::||||||:::::   |:::||
Db   199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAlaAla 218
Qy   775 GGCCTTTATGATAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAGAGCC 834
         |||:::||:::    :::||:::
Db   219 IleLeuTyrPheTyrLeuLeuPheProSer------------SerLeuSer 231
Qy   835 TTCAGCAAAGCTGCCCAGCAACAAAAGAAGCAACAAAAAGAAGCACATGGAGGTGGAGGACAACAGGAGT 894
         |||  |||:::||
Db   232 LeuHisAspArgValAlaValAlaValLysGlyThrTyrGluProGluGluAsp--------- 248
Qy   895 CAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATTGACGTTGAC 954
Db   248 ----------------------------------------- 248
Qy   955 CGGGGAGGAGGAGAAGGGGAAAGACCAATCTGAGAGGATTGTCTTCAGTATGACTA 1014
Db   249 ----------------------------------Trp 249
Qy  1015 GAAGATCGCACTGAAAGCAGACAGAAGACTCCTTAGAACTGTCCTCA 1059
          |||||  |||:::   :::||  |||:::
Db   250 GluAspHisArgGluGluArgLysThrIleGluLeuThrAla 264
```

RESULT 14
AAW94319
ID   AAW94319 standard; protein; 265 AA.
XX
AC   AAW94319;
XX
DT   12-APR-1999 (first entry)
XX
DE   Rat aquaporin-5.
XX
KW   Rat; aquaporin-5; AQP5; AQP1; transmembrane water channel protein;
KW   major intrinsic protein; MIP; CHIP28.
XX
OS   Rattus sp.
XX
PN   US5858702-A.
XX
PD   12-JAN-1999.
XX
PF   24-FEB-1995;   95US-00393996.
XX
PR   13-DEC-1991;   91US-00806273.
PR   17-AUG-1992;   92US-00930168.
XX
PA   (UYJO) UNIV JOHNS HOPKINS.
XX
PI   Agre PC;
XX
DR   WPI; 1999-152100/13.
DR   N-PSDB; AAX04106.
XX
PT   DNA encoding aquaporin-5 water channel protein - useful for producing
PT   recombinant protein.
XX
PS   Claim 1; Col 49-50; 47pp; English.
XX
CC   The present sequence represents a water channel protein called aquaporin-
CC   5 (AQP5). The polynucleotide encoding AQP5 is useful for producing
CC   recombinant AQP5, which can be incorporated into proteoliposomes or cell
CC   membrane vesicles which are able to be used in screening assays for water
CC   channel agonists or antagonists
XX
SQ   Sequence 265 AA;

```
Db          250 GluAspHisArgGluGluArgLysLysThrIleGluLeuThrAla 264

RESULT 15
ADD47948
ID     ADD47948 standard; protein; 265 AA.
XX
AC     ADD47948;
XX
DT     02-DEC-2004  (revised)
DT     29-JAN-2004  (first entry)
XX
DE     Rat Protein AAA66221, SEQ ID NO 13644.
XX
KW     Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW     chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS     Rattus norvegicus.
OS     Unidentified.
XX
PN     WO2003016475-A2.
XX
PD     27-FEB-2003.
XX
PF     14-AUG-2002; 2002WO-US025765.
XX
PR     14-AUG-2001; 2001US-0312147P.
PR     01-NOV-2001; 2001US-0346382P.
PR     26-NOV-2001; 2001US-0333347P.
XX
PA     (GEHO ) GEN HOSPITAL CORP.
PA     (FARB ) BAYER AG.
XX
PI     Woolf C,   D'urso D,   Befort K,   Costigan M;
XX
DR     WPI; 2003-268312/26.
DR     GENBANK; AAA66221.
XX
PT     New composition comprising two or more isolated polypeptides, useful for
PT     preparing a medicament for treating pain in an animal.
XX
PS     Example 1; Page; 1017pp; English.
XX
CC     The invention discloses a composition comprising two or more isolated rat
CC     or human polynucleotides or a polynucleotide which represents a fragment,
CC     derivative or allelic variation of the nucleic acid sequence. Also
CC     claimed are a vector comprising the novel polynucleotide, a host cell
CC     comprising the vector, a method for identifying a nucleotide sequence
CC     which is differentially regulated in an animal subjected to pain and a
CC     kit to perform the method, an array, a method for identifying an agent
CC     that increases or decreases the expression of the polynucleotide sequence
CC     that is differentially expressed in neuronal tissue of a first animal
CC     subjected to pain, a method for identifying a compound which regulates
CC     the expression of a polynucleotide sequence which is differentially
CC     expressed in an animal subjected to pain, a method for identifying a
CC     compound that regulates the activity of one or more of the
CC     polynucleotides, a method for producing a pharmaceutical composition, a
CC     method for identifying a compound or small molecule that regulates the
CC     activity in an animal of one or more of the polypeptides given in the
CC     specification, a method for identifying a compound useful in treating
CC     pain and a pharmaceutical composition comprising the one or more
CC     polypeptides or their antibodies. The polynucleotide or the compound that
CC     modulates its activity is useful for preparing a medicament for treating
CC     pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC     injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC     therapy). The sequence presented is a rat protein (described in Table 3
CC     of the specification) which is differentially expressed during pain.
CC     Note: The sequence data for this patent did not form part of the printed
CC     specification, but was obtained in electronic form directly from WIPO at
CC     ftp.wipo.int/pub/published_pct_sequences.
XX
SQ     Sequence 265 AA;
```

```
Alignment Scores:
Pred. NO.:
Score:                  8.5e-52          Length:              265
Percent Similarity:     614.00           Matches:             133
Best Local Similarity:  56.83%           Conservative:        46
Query Match:            42.22%           Mismatches:          82
DB:                     28.15%           Indels:              54
                        2                Gaps:                5

US-10-723-180-1 (1-1152) x AAW94319 (1-265)

Qy    118 AAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCAGTCACAGCGCGGAATTTCTGGCCATGCTT 177
              |||  |||  :  || |||::||||:|||  ||||:::||||||||
Db      3 LysGluValCysSerLeuAlaPheLysAlaValPheAlaGluPheLeuAlaThrLeu 22

Qy    178 ATTTTGTTCTCCCAGCCTGGGATCCACCATCAACTGGGTGAACAGAAAGCCTTTA 237
              |||||||||          ::  :||||
Db     23 IlePheValPhePheGlyLeuGlySerAlaLeuGlyTrp--------ProSerAlaLeu 39

Qy    238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTTCAGCATTGCAACCATGGTGCAG 297
              ||||                :||::::||||||||||||:::||||
Db     40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy    298 TGCTTTGGCCATATCAGCCGGTGGCCACATCAACCCTGGCAGTGACTGTGGCCATGGTGTGC 357
              ||::        |::||||  ||||||||||||||||::::||||
Db     59 AlaLeuGlyProValSerGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuIle 78

Qy    358 ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCC 417
              ||||||:::       :::||||||||||||||||||||||||
Db     79 GlyAsnGlnIleSerLeuLeuArgAlaValPheTyrValAlaAlaGlnLeuValGlyAla 98

Qy    418 ATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTGGGA 477
              ||:||||||||:|          :::  |||  ||||  :::::|||||
Db     99 IleAlaGlyAlaGlyIleLeuTyrLeuValThrProLeuAlaProLeuAsnAlaArgGlyAsnLeuAla 118

Qy    478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCTCCTGGTTGAGTTGATAATC 537
              ||||||||||||||||||          :::::||||:::||||
Db    119 ValAsnAlaLeuAsnAlaAsnThrThrProGlyLysAlaMetValValGluLeuIleLeu 138

Qy    538 ACATTTCAATTGGTGTTTACTATCTTTGCCCAGCTGTGATTCCAAACGGACTGATGTCACT 597
              |||||                    ||||||||||:::||||
Db    139 ThrPheGlnIleLeuAlaLeuCysIlePheSerSerThrAspSerArgArgThrSerProVal 158

Qy    598 GGGTCAATAGCTTTAGCAATTGGATTTCTGTTGCAATTGACATTTATTTGCAATCAAT 657
              ||||||        ||||||||||||||||||:::||||
Db    159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy    658 TATACTGGTGCCAGCATGAATCCCGCCGATCCTTGGACCTGCAGTTATCATGGGAAAT 717
              :::|||          ||||||:::||||||||||||||||
Db    179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198

Qy    718 TGGGAA--AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCCTGT 774
              :::||||||||        |||  :::||||
Db    199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAlaAla 218

Qy    775 GGGCCTTTATGAGTATGTCTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
              ||||||      |||:::
Db    219 IleLeuTyrPheTyrLeuPheProSer-----------SerLeuSer 231

Qy    835 TTCAGCAAAGCTGCCAGCAAACAAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
              |||||||||||            ::||||
Db    232 LeuHisAspArgValAlaValAlaValLysGlyThrTyrGluProGluGluAsp------- 248

Qy    895 CAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATTGACGTTGAC 954

Db    248

Qy    955 CGGGGAGGAGGAGAAGAGGGGAAAAGACCAATCTGGAGAGGTATTGTCTTCAGTATGACTA 1014

Db    249 --------------------Trp 249

Qy   1015 GAAGATCGCACTGAAAAGCAGCAAGACTCCTTAGAACTGTCCTCA 1059
              ||||||     |||::::  |||||  :|||||:::::::
```

Search completed: September 8, 2005, 12:05:59
Job time : 199.5 secs

Pred. No.:   8.5e-52   Length: 265
Score: 614.00   Matches: 133
Percent Similarity: 56.83%   Conservative: 46
Best Local Similarity: 42.22%   Mismatches: 82
Query Match: 28.15%   Indels: 54
DB: 7   Gaps: 5

US-10-723-180-1 (1-1152) x ADD47948 (1-265)

Qy   118 AAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTT 177
          ||||||||| :: |||||:||||||||||||||||||||| |||||||||||||
Db   3 LysGluValCysSerLeuAlaPheLysAlaValPheAlaValPheLeuAlaThrLeu   22

Qy   178 ATTTTTGTTCCTCCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTA 237
          ||||||||| ||||||| |||||| :: :|| |||
Db   23 IlePheValPheGlyLeuGlySerAlaLeuLysTrp--------ProSerAlaLeu   39

Qy   238 CCGGTCGACATGGTTCTCATCTCCCTTGCTTTGACTCAGCATGCAACCATGTGCAG   297
          ||| :::::: |||||||:::: ||||||||| :::|| ||| ||
Db   40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln   58

Qy   298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
          ||| :::|||||||||||| ||||||||||:::||:: ||::::::
Db   59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuIle 78

Qy   358 ACCAGGAAGATCAGCATCCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCC 417
          :::||||||:: :::::||||||||||||||:::||::||||||||||||:::|||
Db   79 GlyAsnGlnIleSerLeuLeuArgAlaValPheTyrValAlaAlaGlnLeuLeuValGlyAla   98

Qy   418 ATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGGCCTGGGA 477
          ||| ::: ||| ||| :::   :::||::::::
Db   99 IleAlaGlyAlaGlyIleLeuTyrTyrLeuAlaProLeuAlaAsnAlaArgGlyAsnLeuAla 118

Qy   478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTCTCGGTTGAGTTGATAATC 537
          ||| :::::: ||| || ||::: ::::||:::||||||||:::
Db   119 ValAsnAlaLeuAsnAsnThrThrProGlyLysAlaMetValValGluLeuLeuLeu   138

Qy   538 ACATTTCAATTGGTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTTCACT 597
          ||||| |||||||:::||||||||||||||||| :::||
Db   139 ThrPheGlnLeuAlaLeuCysIlePheSerSerThrAspSerArgThrSerProVal 158

Qy   598 GGCTCAATAGCTTTAGCAATTGGATTTCTGTTGCAATTGCAATCAAT 657
          ||||||| ||:::||||| |||||||||:::||||||||
Db   159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisValGlyIleTyr   178

Qy   658 TATACTGGTGCCAGCATGAATCCCGCCCAGATCCTTGGACCTGCAGTTATCATGGGAAAT 717
          :::||||| |||||||||||||||||||||:::|| |||
Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198

Qy   718 TGGGAA---AACCATTGGATATATTGGGTGGGCCCATCATAGGAGCTGTCCTCGCTGGT 774
          :::   ::::::||:::|||||||:::||||||:::|:||
Db   199 PheSerProSerHisTrpValGlyThrValGlyProIleValGlyAlaMetLeuAlaAla 218

Qy   775 GGCCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
          |||||||||:::::
Db   219 IleLeuTyrPheTyrLeuPheProSer----------------------SerLeuSer   231

Qy   835 TTCAGCAAAGCTGCCCAGCAAACAAAGGAAGCTACATGGAGGTCGAGGACAACAGGAGT 894
          |||||||:::|| |||||||:::::
Db   232 LeuHisAspArgValAlaValValLysGlyThrTyrGluProGluGluAsp-------- 248

Qy   895 CAGGTAGAGACGGATGACCTGATTCTAAAACCTGAGTGTGCATGTGATTGACGTTGAC 954
          :::||||||
Db   248 --- 248

Qy   955 CGGGGAGAGGAGAGAAAGGGGAAAGACCAATCTCGGAGAGGTATTGTCTTCAGTATGACTA 1014

Db   249 --Trp 249

Qy 1015 GAAGATCGCACTGAAAGCAGACAAGACTCCTTAGAACTGTCCTCA 1059
          |||||| ||| ::::: ::::||||||||||||||:::::
Db   250 GluAspHisArgGluLeuArgLysLysThrLeuThrLeuGluLeuThrAla 264

This Page Blank (uspto)

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:      September 8, 2005, 11:56:43 ; Search time 41 Seconds
                                  (without alignments)
                                  4194.913 Million cell updates/sec

Title:        US-10-723-180-1
Perfect score: 2181
Sequence:     1 ggggcaggcaatgagagctg........attactcagtctaaacaata 1152

Scoring table: BLOSUM62
              Xgapop  10.0 , Xgapext  0.5
              Ygapop  10.0 , Ygapext  0.5
              Fgapop   6.0 , Fgapext  7.0
              Delop    6.0 , Delext   7.0

Searched:     513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:        1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODELframe+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10723180/runat_08092005_123257_29261/app_query.fasta_1.1351
-DB=Issued_Patents_AA -QFMT=fastan -SUFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10723180 @CGN_1_1_46 @runat_08092005_123257_29261 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :    Issued_Patents_AA:*
              1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
              2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
              3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
              4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
              5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
              6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                                     SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1755	80.5	336	4	US-09-949-016-10148	Sequence 10148, A
2	1755	80.5	336	4	US-09-949-016-10149	Sequence 10149, A
3	624.5	28.6	437	4	US-09-949-016-8094	Sequence 8094, Ap
4	614	28.2	265	1	US-08-468-763-19	Sequence 19, Appl
5	614	28.2	265	4	US-08-393-996A-19	Sequence 19, Appl
6	612	28.1	293	4	US-09-949-016-8027	Sequence 8027, Ap
7	603	27.6	118	4	US-09-513-999C-6089	Sequence 6089, Ap
8	594	27.2	271	1	US-08-447-554-4	Sequence 4, Appli
9	594	27.2	271	1	US-08-448-160-4	Sequence 4, Appli
10	561	25.7	269	1	US-08-468-763-17	Sequence 17, Appl
11	561	25.7	269	2	US-08-393-996A-17	Sequence 17, Appl
12	561	25.7	269	4	US-09-949-016-5934	Sequence 5934, Ap
13	561	25.7	281	4	US-09-949-016-8317	Sequence 8317, Ap
14	561	25.7	281	4	US-09-949-016-10471	Sequence 10471, A
15	560	25.7	269	1	US-08-447-554-5	Sequence 5, Appli
16	560	25.7	269	1	US-08-448-160-5	Sequence 5, Appli
17	426.5	19.6	284	3	US-09-372-448A-4	Sequence 4, Appli
18	419	19.2	288	3	US-09-372-422A-12	Sequence 12, Appl
19	418	19.2	292	3	US-09-372-422A-4	Sequence 4, Appli
20	417	19.1	288	3	US-09-372-448A-2	Sequence 2, Appli
21	417	19.1	289	3	US-09-372-422A-2	Sequence 2, Appli
22	415	19.0	289	3	US-09-372-422A-14	Sequence 14, Appl
23	414	19.0	288	3	US-09-372-422A-18	Sequence 18, Appl
24	410	18.8	296	3	US-09-372-422A-20	Sequence 20, Appl
25	406.5	18.6	292	3	US-09-372-422A-10	Sequence 10, Appl
26	398.5	18.3	254	3	US-09-372-422A-34	Sequence 34, Appl
27	395	18.1	249	3	US-09-372-422A-22	Sequence 22, Appl
28	390.5	17.9	288	3	US-09-372-422A-16	Sequence 16, Appl
29	388	17.8	250	2	US-08-654-025-2	Sequence 2, Appli
30	387	17.7	250	1	US-08-234-939-2	Sequence 2, Appli
31	387	17.7	250	2	US-08-558-865-2	Sequence 2, Appli
32	387	17.7	250	2	US-08-654-025-7	Sequence 7, Appli
33	384.5	17.6	272	3	US-09-372-422A-26	Sequence 26, Appl
34	377.5	17.3	257	3	US-09-372-422A-28	Sequence 28, Appl
35	375	17.2	262	3	US-09-372-422A-32	Sequence 32, Appl
36	364	16.7	249	3	US-09-372-448A-6	Sequence 6, Appli
37	358	16.4	281	3	US-09-053-702-2	Sequence 2, Appli
38	353.5	16.2	249	3	US-09-372-422A-30	Sequence 30, Appl
39	352.5	16.2	250	3	US-09-372-422A-24	Sequence 24, Appl
40	334	15.3	247	3	US-09-372-422A-48	Sequence 48, Appl
41	332.5	15.2	294	3	US-09-372-422A-40	Sequence 40, Appl
42	320.5	14.7	295	3	US-09-372-422A-38	Sequence 38, Appl
43	311.5	14.3	263	4	US-09-610-906-12	Sequence 12, Appl
44	306.5	14.1	262	4	US-09-976-594-347	Sequence 347, App
45	306.5	14.1	443	4	US-09-610-906-1	Sequence 1, Appli

```
                                    ALIGNMENTS

RESULT 1
US-09-949-016-10148
; Sequence 10148, Application US/0994916
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10148
;  LENGTH: 336
;  TYPE: PRT
;  ORGANISM: Human
US-09-949-016-10148

Alignment Scores:
Pred. No.:                     2.89e-180
Score:                         1755.00                 Length:          336
Percent Similarity:            100.00%                 Matches:         336
Best Local Similarity:         100.00%                 Conservative:    0
Query Match:                   80.47%                  Mismatches:      0
DB:                            4                       Indels:          0
                                                       Gaps:            0

US-10-723-180-1 (1-1152) x US-09-949-016-10148 (1-336)

Qy      1 GGGGCAGGCAATGAGAGCTGCACTCTGGCTCGGCTGGGAAGGCATGAGGTGACAGACCCACAGCA 60
```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10149
;   LENGTH: 336
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10149

Alignment Scores:
Pred. No.:   2.89e-180   Length: 336
Score: 1755.00   Matches: 336
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 80.47%   Indels: 0
DB: Gaps: 4

US-10-723-180-1 (1-1152) x US-09-949-016-10149 (1-336)

Qy   1 GGGCAGGCAATGAGAGCTGCACTCTGGCTGGGGAAGGCATGAGTGACAGACCCACAGCA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GlyAlaGlyAsnGluSerCysThrLeuAlaGlyGluGlyMetSerAspArgProThrAla   20

Qy   61 AGGCGGTGGGTAAGTGTGACCTTTGTGTACCAGAGAGAACATCATGTGGCTTTCAAA   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 ArgArgTrpGlyLysCysThrArgGluAsnIleMetValAlaPheLys   40

Qy 121 GGGGTCTGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTTATT   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 GlyValTrpThrGlnAlaPheThrLysAlaValThrAlaGluPheLeuAlaMetLeuIle   60

Qy 181 TTTGTTCTCCTCAGCCTGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCG   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 PheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlyThrGluLysProLeuPro   80

Qy 241 GTCGACATGGTTCCATCCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGC   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 ValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetValGlnCys   100

Qy 301 TTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGCACC   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 PheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThr   120

Qy 361 AGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCCAGTGCCTGGGGGCCATC   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 ArgLysIleSerIleAlaLysSerValPheTyrIleAlaSerAlaGlnCysLeuGlyAlaIle   140

Qy 421 ATTGGAGCAGGAATCCTCATCTGCAGTGGTGGGAGGGTCTGGGAGTC   480
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 IleGlyAlaGlyIleLeuIleTyrLeuValThrProProSerValGlyLeuGlyVal   160

Qy 481 ACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCCTGGTTGAGTTGAATAACACA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 ThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuLeuValGluLeuIleIleThr   180

Qy 541 TTTCAATTGGTGTTACATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGC   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 PheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgThrAspValThrGly   200

Qy 601 TCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGACATTATTGGACAATCAATTAT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 SerIleAlaLeuAlaIleAlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTyr   220

RESULT 2
US-09-949-016-10149
; Sequence 10149, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

Db   1 GlyAlaGlyAsnGluSerCysThrLeuAlaGlyGluGlyMetSerAspArgProThrAla   20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1 GlyAlaGlyAsnGluSerCysThrLeuAlaGlyGluGlyMetSerAspArgProThrAla   20

Qy 61 AGGCGGTGGGTAAGTGTGACCTTTGTGTACCAGAGAGAACATCATGTGGCTTTCAAA   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 21 ArgArgTrpGlyLysCysThrArgGluAsnIleMetValAlaPheLys   40

Qy 121 GGGGTCTGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTTATT   180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41 GlyValTrpThrGlnAlaPheThrLysAlaValThrAlaGluPheLeuAlaMetLeuIle   60

Qy 181 TTTGTTCTCCTCAGCCTGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCG   240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 PheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlyThrGluLysProLeuPro   80

Qy 241 GTCGACATGGTTCCATCCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGC   300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 81 ValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetValGlnCys   100

Qy 301 TTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGCACC   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 PheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThr   120

Qy 361 AGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCCAGTGCCTGGGGGCCATC   420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 ArgLysIleSerIleAlaLysSerValPheTyrIleAlaSerAlaGlnCysLeuGlyAlaIle   140

Qy 421 ATTGGAGCAGGAATCCTCATCTGCAGTGGTGGGAGGGTCTGGGAGTC   480
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 IleGlyAlaGlyIleLeuIleTyrLeuValThrProProSerValGlyLeuGlyVal   160

Qy 481 ACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCCTGGTTGAGTTGAATAACACA   540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 ThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuLeuValGluLeuIleIleThr   180

Qy 541 TTTCAATTGGTGTTACATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGC   600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 PheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgThrAspValThrGly   200

Qy 601 TCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGACATTATTGGACAATCAATTAT   660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 SerIleAlaLeuAlaIleAlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTyr   220

Qy 661 ACTGGTGCCAGCATGAATCCCGCCCGATCCTTTGGACCTGCCAGTTATCATGGGAAATTGG   720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 ThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTrp   240

Qy 721 GAAAACCATTGGATATATGGGTGGGCCCATCATAGGAGCTGTCCTCGCTGGTGGCCTT   780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaValLeuAlaGlyGlyLeu   260

Qy 781 TATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGAGCCTTCAGC   840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 261 TyrGluTyrValPheCysProAspValGluPheLysArgArgPheLysGluGluAlaPheSer   280

Qy 841 AAAGCTGCCCAGCAAACAAAAGGAGAGTACATGGAGGTGGAGGACAACAGGAGTCAGTA   900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 281 LysAlaAlaGlnGlnThrLysGluGluSerTyrMetGluValGluAspAsnArgSerGlnVal   300

Qy 901 GAGACGGATGACCTGATTCTAAAACCTGGAGTGGTCATGTGGTGCATGGTCACGTTGACCGTGACCGGGA   960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GluThrAspAspLeuIleLeuLysProGlyValValHisValAspProSerValAspArgArgGly   320

Qy 961 GAGGAGAAGGGGAAAAGACCAATCTGAGAGGGTATTGTTCTTCCAGTA   1008
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 321 GluGluLysLysGlyLysLysAspGlnSerGlyGlyValLeuSerSerVal   336

```
Qy  661 ACTGGTGCCAGCATGAATCCCGCCGATCCTTGGACCTGCAGTTATCATGGGAAATTGG 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 ThrGlyAlaSerMetAsnProAlaAspProTrpThrCysSerTyrHisMetGlyAsnTrp 240

Qy  721 GAAAACCATTGGATATATGGGTTGGCCCCATCATAGGAGCTGTCCTGCTGGTGGCCTT 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaValLeuLeuValAlaLeu 260

Qy  781 TATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGCCTTCAGC 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261 TyrGluTyrValPheCysProAspValGluPheLysArgArgPheLysGluAlaPheSer 280

Qy  841 AAAGTGCCCAGCAAACAAAGGAGCTACATGGAGGTGGAGGACAACAGGAGTCAGGTA 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  281 LysAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAspAsnArgSerGlnVal 300

Qy  901 GAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATTGACGTTGACCGGGGA 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GluThrAspAspLeuIleLeuLysProGlyValValHisValIleAspValAspArgGly 320

Qy  961 GAGGAGAAGAAGGGGAAAGACCAATCTGGAGAGGTATTGTGTCTTCAGTA 1008
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  321 GluGluLysLysGlyLysAspGlnSerGlySerGlyValValLeuSerVal 336
```

RESULT 3
US-09-949-016-8094
; Sequence 8094, Application US/0994016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8094
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8094

```
Alignment Scores:
Pred. No.:
Score:                   2.85e-58        Length:        437
                         624.50          Matches:       140
Percent Similarity:      62.15%          Conservative:  39
Best Local Similarity:   48.61%          Mismatches:    88
Query Match:             28.63%          Indels:        21
DB:                      4               Gaps:          7
```

US-10-723-180-1 (1-1152) x US-09-949-016-8094 (1-437)

```
Qy   52 CCCACA--GCAAGGCGGTGGGGTAAGTGT-----GGACCTTTGTGTACCAGAGAGAAC 102
        ||||||    |||   |||  ||| |||        |||  :::
Db  156 ProThrGlyAlaProArgArgGlySerCysArgArgAlaProAlaAlaThrMetLys--- 174

Qy  103 ATCATGGTGCGCTTTCAAAGGGGTCTGGAGCTTCAAGCTTTCTGGAAAGCAGTCACGCGGAA 162
        |||   :::  |||||||||||||||  |||:::
Db  175 --------LysGluValCysSerValAlaPheLeuLysLysAlaValPheAlaGlu 189

Qy  163 TTTCTGGCCATGCTTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGA 222
        |||||||||||||||||||||||||||||||||||||||||||:::
Db  190 PheLeuAlaThrLeuIlePheValPhePheLysGlySerAlaLeuLysTrp------ 207

Qy  223 ACAGAAAAGCCTTTACCGGTGACATGGTTCTCATCTCCCTTTGCTTTGGACTTCAGCATT 282
        |||||||||||||||   |||||||   |||:::||   |||||:||||||:::|||
```

```
Db  208 ---ProSerAlaLeuProThr---IleLeuGlnIleAlaLeuAlaLeuAlaPheGlyLeuAlaIle 225
Qy  283 GCAACCATGGTGCAGTGCTTGGCCATATCACCGGTGGCCACATCACCCTGCCAGTGACT 342
        |||||||||:::||     |||    :::||||||||  |||||||||||:|||:|||
Db  226 GlyThrLeuAlaGlnAlaLeuGlyProValSerGlyValHisIleAsnProAlaIleThr 245
Qy  343 GTGGCCATGGTGTGCACCAGGAAGATCAGCATGCCAAGTCGTCTTCTACATCGCAGCC 402
        ::|||:||||||     :::||||     ::::|| ||||||:|||||||:::||||||
Db  246 LeuAlaLeuLeuValGlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAla 265
Qy  403 CAGTGCCTGGGGCCATCATTGGAGCAGGAATCCTATCGGTCACACCTTCCAGTGTG 462
        |||||||||:||    |||||||||||||||||:|||:||||||||||:|||||||
Db  266 GlnLeuValGlyAlaIleAlaGlyAlaGlyIleLeuLeuTyrGlyValAlaProLeuAsnAla 285
Qy  463 GTGGAGGGCCTGGAGTCACCATGGTCACCACGGCTGGTCATGGTCTCCTG 522
        |||||||||||     ::|||:||||||||||||||:::|||
Db  286 ArgGlyAsnLeuAlaValAsnAlaLeuAsnAsnThrThrGlnAlaMetVal 305
Qy  523 GTTGAGTTGATAATCACATTTCAATTGGTGTTTACTATCTTTGCCAGTGTGATTCCAAA 582
        |||||||||:::||||||||||||||     ::|||||||||||||||||||||
Db  306 ValGluLeuLeuThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArg 325
Qy  583 CGGACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGACAT 642
        |||||||||     ::|||||||||:::||||||||||||||:|:|||||||:|||||
Db  326 ArgThrSerProValGlySerProAlaLeuSerIleGluGlyLeuSerValThrLeuGlyHis 345
Qy  643 TTATTTGCAATCAATTATACTGGTGCCAGCATGATGAATCCCGCCGATCCTTGGACCTGCA 702
        |||    |||||||||||||||||||||||||||:::||||||||||||:::||||||
Db  346 LeuValGlyIleTyrPheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAla 365
Qy  703 GTTATCATGGGAAATTGGGAA---AACCATTGGATATATGGGTTGGCCCCATCATAGGA 759
        |||||||||||:::|| ||    |||||||||||||||||:::|||||||:||||||||
Db  366 ValValMetAsnArgPheSerProAlaHisTrpValPheTrpValGlyProIleValGly 385
Qy  760 GCTGTTCCTGCTGGTGGCCTTTATGAGTAGTGTCTTCTGTCCAGATGTTGAATTCAAACGT 819
        |||||||||||||||:::||    ::||||:||||||||||||||||||:||||||||
Db  386 AlaValLeuAlaLeuAlaIleAlaIleLeuTyrPheTyrLeuLeuPheProAsn------ 400
Qy  820 CGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAACAAAGAGCTACATGGAGGTG 879
        |||||||||||||:::|||||||||||||||||||||||||:||:|||||:|||||
Db  401 ------SerLeuSerGluArgValAlaIleIleLysLysGlyThrTyrGluProAsp 418
Qy  880 GAGGACAACAGGAGTCAGGTAGAG 903
        ||||||||||||||||:|||
Db  419 GluAspTrpGluGluGlnArgGlu 426
```

RESULT 4
US-08-468-763-19
; Sequence 19, Application US/08468763
; Patent No. 5741671
; GENERAL INFORMATION:
; APPLICANT: Agre, Peter C.
; TITLE OF INVENTION: Isolation, Cloning and Expression of
; TITLE OF INVENTION: Transmembrane Water Channel Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,763
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,996

```
Qy    835  TTCAGCAAAGTGCCCAGCAACAAACAAAGGAAGAAGCTACATGGAGGTGGAGGACAACAGGAGT  894
                |||||  ||  :::||||
Db    232  LeuHisAspArgValAlaValValGlyThrTyrGluProGluGluAsp---------  248

Qy    895  CAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATTGATTGACGTTGAC  954
                -----------------------------------------------------------  ----
Db    248  ------------------------------------------------------------- 248

Qy    955  CGGGGAGGAGAGAAGGGGAAAGACCAATCTGGAGAGTATTGTCTTCAGTATGACTA 1014
                ---------------------------------------------------------
Db    249  ------------------------------------------------------Trp 249

Qy   1015  GAAGATCGCACTGAAAGCACAAGACAAGACTCCTTAGAACTGTCCTCA 1059
                |||||  ||||:  :::||||:||||||:||:
Db    250  GluAspHisArgGluGluGluArgLysLysThrIleGluLeuThrAla 264


RESULT 5
US-08-393-996A-19
; Sequence 19, Application US/0839396A
; Patent No. 5858702
; GENERAL INFORMATION:
;    APPLICANT: Agre, Peter C.
;    TITLE OF INVENTION: Isolation, Cloning and Expression of
;    TITLE OF INVENTION: Transmembrane Water Channel Proteins
;    NUMBER OF SEQUENCES: 19
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Banner & Allegretti
;       STREET: 1001 G Street, N.W.
;       CITY: Washington, D.C.
;       STATE: D.C.
;       COUNTRY: US
;       ZIP: 20001
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/393,996A
;       FILING DATE: 24-FEB-1995
;       CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Posorske, Laurence H.
;       REGISTRATION NUMBER: 34,698
;       REFERENCE/DOCKET NUMBER: 1107.48633
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 202 508-9100
;       TELEFAX: 202 508-9299
;    INFORMATION FOR SEQ ID NO: 19:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: 265 amino acids
;          TYPE: amino acid
;          TOPOLOGY: linear
;          MOLECULE TYPE: protein
; US-08-393-996A-19

Alignment Scores:
Pred. No.:               19:
Score:                   2.93e-57          Length:          265
Percent Similarity:      614.00            Matches:         133
Best Local Similarity:   56.83%            Conservative:    46
Query Match:             42.22%            Mismatches:      82
DB:                      28.15%            Indels:          54
                         2                 Gaps:            5

US-10-723-180-1 (1-1152) x US-08-393-996A-19 (1-265)

Qy    118  AAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTT 177
                |||  ||   :::  |||||||| :|||:|||||||||||||||
Db      3  LysGluValCysSerLeuAlaPheLysAlaValPheAlaGluPheLeuAlaThrLeu 22

Qy    178  ATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACCTGGGGTGGAACAGAAAGCCTTTA 237
                ||||||||||||||       |||
Db     23  IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp---------ProSerAlaLeu 39
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8027
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8027

Alignment Scores:
Pred. No.:                              Length:        293
Score:                    5.09e-57      Matches:       114
Percent Similarity:       65.27%        Conservative:   57
Best Local Similarity:    43.51%        Mismatches:     77
Query Match:              28.06%        Indels:         14
DB:                                     Gaps:            3
```

US-10-723-180-1 (1-1152) x US-09-949-016-8027 (1-293)

```
Qy    124 GTCTGG-------ACTCAAGCTTCTGAAAGACAGTCACAGCGGAATTTCTGCCCATG  174
          :::|||        ::||||||||||   :||||| :| ::||||
Db     31 MetTrpGluLeuArgSerAlaSerPheTrpArgAlaIlePheAlaGluPheAlaThr   50

Qy    175 CTTATTTTGTTCTCCTGCAGGATCCAACTCACCATCAACTGGGGTGGAACAGAAAAGCCT  234
          ||||                    :::|||||||||| ::::::              --Pro   66
Db     51 LeuPheTyrValPhePheGlyLeuGlySerSerLeuArgTrpAla--------------Pro

Qy    235 TTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTGCAACCATGGTG  294
          ||||||||| ||||||||||||||||||||           ||||||||| :||||||
Db     67 GlyProLeuHisValLeuGlnValAlaLeuAlaMetAlaPheGlyLeuAlaLeuAlaThrLeuVal   86

Qy    295 CAGTGCTTGGCCATATCAGCGGTGGCCACCATCAACCCTGCCAGTGACTGTGGCCATGGTG  354
          ||||||||||||||||||||||||||||||:::|||||||||||||||||         :::
Db     87 GlnSerValGlyHisIleSerGlyAlaHisValAsnProAlaValThrPheAlaPheLeu  106

Qy    355 TGCACCAGGAAGATCAGCATGCCCAAGTCGTCTTCTACATCGCCAGTCGGGG  414
          :::::      :::::       :::|||||||||||||||||||||||||||
Db    107 ValGlySerGlnMetSerLeuLeuArgAlaPheCysTyrMetAlaLeuLeuGly  126

Qy    415 GCCATCATTGGAGCAGGAGAATCCTTATCTGGTCACACCTCCCAGTGTGGGTGGGAGGCCTG  474
          ||||||:::::   :::::|||||| :::|||::||||||||||||||      :::
Db    127 AlaValAlaGlyAlaValAlaValLeuTyrSerValThrProProAlaValArgGlyAsnLeu  146

Qy    475 GGAGTGCACCATGGTTCATGGAAATCTTACCGGCTGGTCATGGTGTCCTCGGTTGAGTTGATA  534
          |||||||||:::        ::::::|||
Db    147 AlaLeuAsnThrLeuHisProAlaValSerValGlyGlnAlaThrThrValGluIlePhe  166

Qy    535 ATCACATTTCAATTGTGTTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTC  594
          :::|||   ::|||               :::|||   :::||
Db    167 LeuThrLeuGlnPheValLeuCysIlePheAlaAlaThrTyrAspGluArgArgAsnGlyGln  186

Qy    595 ACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGACATTATTGCAATC  654
          |||   :::      |||||||||||||:::||||||||||||||||||||||
Db    187 LeuGlySerValAlaLeuAlaValGlyPheSerLeuAlaLeuGlyHisLeuPheGlyMet  206

Qy    655 AATTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATCGGA  714
          |||||||||||||||:::|||||||||||||||::::::||||||||||||||||
Db    207 TyrTyrThrGlyAlaGlyMetAsnProAlaIleProAlaIleLeuThrLeuGlnThrGly  226

Qy    715 AATTGGGAAAACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCTGGT  774
          ||||:::: |||||||||||||||||||||:::||||||||||||||||
Db    227 AsnPheThrAsnHisTrpValTyrTrpValGlyLeuGlyProIleIleGlyAlaValLeuGlySer  246

Qy    775 GGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGCC  834
          ||||||||||||||||||||||||||:::          -----LysSer   259
Db    247 LeuLeuTyrAspPheLeuGlyLeuPheProArgLeu----------------LysSer

Qy    835 TTCAGCAAAGCTGCCCAGCAAACAAAAGAAGTCACATGAGGTGGAGGACCAACAGGAGT  894
          ||||| ::|                :::   ||||||
Db    260 IleSerGluArgLeuSerValLeuLysProAlaLysProAspValSerAsnGlyGlnPro  279

Qy    895 CAGGTA                                                         900
```

```
Db        280 GluVal 281
              :::|||
          :::|||
```

```
Db        101 ThrValAlaMetValCysThrArg***IleSerIleAlaLysSerValPheTyr 118
              ||||||||||||||||||||   |||||||||||||||||||||||||||||||||
```

RESULT 7
US-09-513-999C-6089
; Sequence 6089, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Duclert, A.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;   Patent No. 6783961
;   FILE REFERENCE: 59.US2.REG
;   CURRENT APPLICATION NUMBER: US/09/513,999C
;   CURRENT FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/122,487
;   PRIOR FILING DATE: 1999-02-26
;   NUMBER OF SEQ ID NOS: 36681
;   SOFTWARE: Patent.pm
; SEQ ID NO 6089
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 57
;   OTHER INFORMATION: Xaa=Ile or Met
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 58
;   OTHER INFORMATION: Xaa=Asp or Asn
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 75
;   OTHER INFORMATION: Xaa=Leu or Arg
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 109
;   OTHER INFORMATION: Xaa=Glu or Lys
US-09-513-999C-6089

Alignment Scores:
Pred. No.:   2.86e-56   Length: 118
Score: 603.00   Matches: 114
Percent Similarity: 96.61%   Conservative: 0
Best Local Similarity: 96.61%   Mismatches: 4
Query Match: 27.65%   Indels: 0
DB: 4   Gaps: 0

US-10-723-180-1 (1-1152) x US-09-513-999C-6089 (1-118)

Qy   40 ATGAGTGACAGACCCACAGCCAAGCGCGGTGGGTAAGTGTGGGACCTTTGTGTACCAGAGAG 99
         |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1 MetSerAspArgProThrAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20

Qy   100 AACACTCATGGTGGCTTCAAAGGGGTCTGGACTTCAAGCTTCTGGAAAGCAGTCACAGCG 159
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40

Qy   160 GAATTTCTGCCATGCTTATTTTGTTCTCCAGCCTGGATCCACCATCAACTGGGGT 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThr*****TrpGly 60

Qy   220 GGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTTGGACTCAGC 279
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GlyThrGluLysProLeuProValAspMetValLeuIleSer***CysPheGlyLeuSer 80

Qy   280 ATTGCAACCATGGTGCAGTGCTTTGGCCCATATCAGCGGTGGCCACATCAACCCTGCAGTG 339
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyLyGlyHisIleAsnProAlaVal 100

Qy   340 ACTGTGGCCCATGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTAC 393
```

RESULT 8
US-08-447-554-4
; Sequence 4, Application US/08447554
; Patent No. 5661003
; GENERAL INFORMATION:
;   APPLICANT: FUSHIMI, KIYOHIDE
;   APPLICANT: UCHIDA, SHINICHI
;   APPLICANT: SASAKI, SEI
;   APPLICANT: MARUMO, FUMIAKI
;   TITLE OF INVENTION: WATER CHANNEL
;   NUMBER OP SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Morrison & Foerster
;     STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
;     CITY: Washington, DC
;     COUNTRY: USA
;     ZIP: 20006-1812
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/447,554
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/126,365
;     FILING DATE: 24-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Murashige, Kate H.
;     REGISTRATION NUMBER: 29,959
;     REFERENCE/DOCKET NUMBER: 5100-0003.00
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202) 887-1500
;     TELEFAX: (202) 887-0763
;     TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 271 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-447-554-4

Alignment Scores:
Pred. No.:                                4.26e-55      Length:          271
Score:                                    594.00        Matches:         111
Percent Similarity:                       66.67%        Conservative:    39
Best Local Similarity:                    49.33%        Mismatches:      71
Query Match:                              27.24%        Indels:          4
DB:                                       1             Gaps:            1

US-10-723-180-1 (1-1152) x US-08-447-554-4 (1-271)

Qy   136 GCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCTCCTCAGC 195
         ||||   ::|||||||| |  |:|||   ||:::|||||||||||||||
Db     8 AlaPheSerArgAlaValLeuAlaGluPheLeuAlaThrLeuLeuPheValPhePheGly 27

Qy   196 CTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTC 255
         ||||||||  ::  ::  |||   |:   |||||||--------ProProSerValLeuGln 43
Db    28 LeuGlySerAlaLeuGlnTrpAlaSerAsr-------

Qy   256 ATCTCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCCATATCAGC 315
         ||:|||:::     |||  |||  ||||   ||:  |:|||
Db    44 IleAlaValAlaPheGlyLeuGlyIleGlyIleLeuLeuGluValGlnAlaLeuGlyHisValSer 63

Qy   316 GGTGGCCACATCAACCCTGCAGTGACTGTGGCCCATGTGTGCACCAGGAAGATCAGCATC 375
         ||||||||||||||||:   :::
Db    64 GlyAlaHisIleAsnProAlaValThrValAlaCysLeuValGlyCysHisValSerPhe 83
```

```
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 271 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-448-160-4

Alignment Scores:
Pred. No.:                           Length:              271
Score:                4.26e-55       Matches:             111
Percent Similarity:   594.00         Conservative:         39
Best Local Similarity: 66.67%        Mismatches:           71
Query Match:          49.33%         Indels:                4
DB:                   27.24%         Gaps:                  1
                      1

US-10-723-180-1 (1-1152) x US-08-448-160-4 (1-271)

Qy  136  GCTTTCTGGAAAGCACTCACGCGGAATTTCTGGCCATGCTTTATTTTGTCTCCTCAGC  195
              :::||||||||||   ||||||||||||  ||||:::|||
Db    8  AlaPheSerArgAlaValLeuAlaGlnLeuLeuGlyLeuAlaThrLeuPheValPhePheGly  27

Qy  196  CTGGGATCCACCATCAACTCGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTC  255
         ||||||||||   ::: |||      :::
Db   28  LeuGlySerAlaLeuGlnTrpAlaSerSer--------ProProSerValLeuGln  43

Qy  256  ATCTCCCTTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCATATCAGC  315
         ||||||||   ||||||  :::|||       ::||||
Db   44  IleAlaValAlaAlaPheGlyLeuGlyIleGlyIleGlnAlaLeuGlnHisValSer  63

Qy  316  GGTGGCCACATCAACCCTGCAGTGACTGCTGGTGTGCACCAGGAGACAGCAGCATC  375
         ||||||||   |||||||       :::
Db   64  GlyAlaHisIleAsnProAlaValThrAlaCysLeuValGlyCysHisValSerPhe  83

Qy  376  GCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCCATCATTGGAGCAGGAATC  435
              :::||||||||||   ||||||||||||  ||||:::|||
Db   84  LeuArgAlaAlaPheTyrValAlaAlaGlnLeuLeuGlyLeuAlaValAlaGlyAlaAlaIle  103

Qy  436  CTCTATCTGGTCACACCTCCAGTGTGGTGGAGGCCTGGAGTCACCATGGTTCATGGA  495
         |||:::   ::||||||||    :::   |||
Db  104  LeuHisGluIleThrProValGluIleArgGlyAspLeuAlaValAlaAsnAlaLeuHisAsn  123

Qy  496  AATCTTACCGCTGGTCATGGTCTCCTGGTTGAGTTGATAATCACATTTCAATTGGTGTTT  555
         |||  |||||||||||     :::||    ||||||||
Db  124  AsnAlaThrAlaGlyGlnAlaValThrValGluLeuPheLeuThrMetGlnLeuValLeu  143

Qy  556  ACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCAATAGCTTTAGCA  615
         |||||||||||   :::|||       |||
Db  144  CysIlePheAlaSerThrAspGluArgArgGlyAspAsnLeuGlySerProAlaLeuSer  163

Qy  616  ATTGGATTTTCTGTTGCAATTGGACATTATTTGCAATCAATTATACTGGTGCCAGCATG  675
         ||||||||||||   :::|||||||    :::|||
Db  164  IleGlyPheSerValThrLeuGlyHisLeuLeuGlyIleTyrPheThrGlyCysSerMet  183

Qy  676  AATCCCGCCCGATCCTTTGGACCTGCAGTTATCATGGGAAATTGGGAAAAACCATTGGATA  735
         |||||||||   ||||||  :::||||       :::||||
Db  184  AsnProAlaArgSerLeuAlaProAlaValAlaThrGlyLysPheAspAspHisTrpVal  203

Qy  736  TATTGGGTTGGGCCCATCATAGGAGAGCTGTCCTCCGCTGGTGCCTTTATGAGTATGTCTTC  795
         |||:::||||||||   :::|||||||       :::|||
Db  204  PheTrpIleGlyProLeuValGlyAlaIleIleGlySerLeuLeuTyrLeuLeu  223

Qy  796  TGTCCAGATGTTGAA  810
         ||| :::
Db  224  PheProSerAlaLys  228
```

```
RESULT 10
US-08-468-763-17
;  Sequence 17, Application US/08468763
;  Patent No. 5741671
;  GENERAL INFORMATION:
;  APPLICANT: Agre, Peter C.
;  TITLE OF INVENTION:  Isolation, Cloning and Expression of
;  TITLE OF INVENTION:  Transmembrane Water Channel Proteins
```

```
RESULT 9
US-08-448-160-4
;  Sequence 4, Application US/08448160
;  Patent No. 5785986
;  GENERAL INFORMATION:
;  APPLICANT: FUSHIMI, KIYOHIDE
;  APPLICANT: UCHIDA, SHINICHI
;  APPLICANT: SASAKI, SEI
;  APPLICANT: MARUMO, FUMIAKI
;  TITLE OF INVENTION:  WATER CHANNEL
;  NUMBER OF SEQUENCES: 8
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Morrison & Foerster
;  STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
;  CITY: Washington, DC
;  COUNTRY: USA
;  ZIP: 20006-1812
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/448,160
;  FILING DATE:
;  CLASSIFICATION: 424
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: US/08/447,554
;  FILING DATE:
;  APPLICATION NUMBER: US 08/126,365
;  FILING DATE: 24-SEP-1993
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Murashige, Kate H.
;  REGISTRATION NUMBER: 29,959
;  REFERENCE/DOCKET NUMBER: 5100-0003.00
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (202) 887-1500
;  TELEFAX: (202) 887-0763
;  TELEX: 90-4030
;  INFORMATION FOR SEQ ID NO:  4:
```

```
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Banner & Allegretti
;     STREET: 1001 G Street, N.W.
;     CITY: Washington, D.C.
;     STATE: D.C.
;     COUNTRY: US
;     ZIP: 20001
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/468,763
;     FILING DATE: 06-JUN-1995
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/393,996
;     FILING DATE: 24-FEB-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Posorske, Laurence H.
;     REGISTRATION NUMBER: 34,698
;     REFERENCE/DOCKET NUMBER: 1107.48633
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202 508-9100
;     TELEFAX: 202 508-9299
;   INFORMATION FOR SEQ ID NO: 17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 269 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
US-08-468-763-17

Alignment Scores:
Pred. No.:               1.54e-51         Length:           269
Score:                   561.00           Matches:          118
Percent Similarity:      43.70%           Conservative:      46
Best Local Similarity:   43.70%           Mismatches:        88
Query Match:             25.72%           Indels:            18
DB:                      1                Gaps:               4

US-10-723-180-1 (1-1152) x US-08-468-763-17 (1-269)

Qy  139 TTCTGGAAAGCAGTCACAGCGGAATTTCGGCCATGCTGTATTTTGTTCTCCCTCAGCCTG 198
            |||||::::||||||  :: || ||||||       ::::|||||  ::|||||:||||
Db   10 PheTrpArgAlaValValAlaGluPheLeuAlaThrLeuPheLeuValPheIleSerIle  29

Qy  199 GGATCCACCATCAACTGGGGT--------GGAACAGAAAAGCCTTTACCGGTCGACATG 249
            ||||||  ||| :::::         ||| ::: ||
Db   30 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn  49

Qy  250 GTTCTCATCTCCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
            |||||||::||| ||| ||||||||||||||||::::
Db   50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis  69

Qy  310 ATCAGCGGTGGCCACATCAACCCCGCAGTGACTGTGGCCATGGTGTGCACCAGGAAGATC 369
            ||||||:||||||||||||||||:  ::::  ::: :||||||
Db   70 IleSerGlyAlaHisIleLeuAsnProAlaValThrLeuGlyLeuLeuLeuSerCysGlnIle  89

Qy  370 AGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTCGGGGGCCATCATTGGAGCA 429
            ||||| ::::::|||  ::: ::||||||:::||||||||||:
Db   90 SerIlePheArgAlaLeuMetTyrIleIleIleIleIleGluIleThrAlaValAlaThr 109

Qy  430 GGAATCCCTATCGTCACACCTCCCAGTGTGGGAGGGCCTGGAGTCACCATGGTT 489
            |||||| :::||| ||  :::::||||||||
Db  110 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 129

Qy  490 CATGGAAATCTTACCGCTGGTCATGGTGTCTCCTGTTGATAATCACCATTTCAATTG 549
            ::: ||:|||||  :::|||||||| ||| ||||||
Db  130 AlaAspGlyValAsnSerGlyValGlnGlyLeuGlyLeuGluIleIleGlyThrLeuGlnLeu 149
```

```
Qy  550 GTGTTACTATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGCTCAATAGCT 609
            ||| ::: ||||||||||||
Db  150 ValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlyGlySerAlaPro 169

Qy  610 TTAGCAATGGATTTCTGTGCAATTGGACATTTATTTGCAATCAATTACTGGTGCC 669
            |||||||||::: ||||||   :::||||
Db  170 LeuAlaIleIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrGlyCys 189

Qy  670 AGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGAAAATTGGGAAAACCAT 729
            ::: |||||||||| ||||||:::
Db  190 GlyIleLeuAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209

Qy  730 TGGATATATTGGGTGGGCCATCATACAGGAGCTGTCCTCCGCTGGTGGCCTTTATGAGTAT 789
            ||||:::::||||||||||||||:::::::
Db  210 TrpIlePheTrpValGlyProPheIleGlyGlyAlaLeuAlaValLeuIleTyrAspPhe 229

Qy  790 GTCTTCTGTCCAGATGTTGAATTCAAACGTGTTTAAAGAAGCCTTCAGCAAAGCTGCC 849
            ||| ::: ::
Db  230 IleLeuAlaPro-------------ArgSerSerAspLeuThrAspArgValLys 243

Qy  850 CAGCAAACAAAAGGAAGCTAC---------ATGGAGGTGGAGGACAACAGGAGTCAG 897
            ||| ||  ::||||||
Db  244 ValTrpThrSerGlyValGlnValGluTyrAspLeuAspAlaAspIleAsnSerArg 263

Qy  898 GTAGAGACGGATGACCTGATTCTAAAACCT 927
            ::: |||||||
Db  264 ValGlu----------MetLysPro 268
```

```
RESULT 11
US-08-393-996A-17
; Sequence 17, Application US/08393996A
; Patent No. 5858702
; GENERAL INFORMATION:
;   APPLICANT: Agre, Peter C.
;   TITLE OF INVENTION: Isolation, Cloning and Expression of
;   TITLE OF INVENTION: Transmembrane Water Channel Proteins
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Banner & Allegretti
;     STREET: 1001 G Street, N.W.
;     CITY: Washington, D.C.
;     STATE: D.C.
;     COUNTRY: US
;     ZIP: 20001
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/393,996A
;     FILING DATE: 24-FEB-1995
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Posorske, Laurence H.
;     REGISTRATION NUMBER: 34,698
;     REFERENCE/DOCKET NUMBER: 1107.48633
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202 508-9100
;     TELEFAX: 202 508-9299
;   INFORMATION FOR SEQ ID NO: 17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 269 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
US-08-393-996A-17

Alignment Scores:
Pred. No.:               1.54e-51         Length:           269
Score:                   561.00           Matches:          118
Percent Similarity:      60.74%           Conservative:      46
Best Local Similarity:   43.70%           Mismatches:        88
```

Query Match: 25.72%   Indels: 18
DB: 2   Gaps: 4

US-10-723-180-1 (1-1152) x US-08-393-996A-17 (1-269)

```
Qy  139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCCTCCAGCCTG 198
        ||||||||::||||| ||||||||||||||||||||       ::::||||:::
Db   10 PheTrpArgAlaValValAlaGluPheLeuAlaThrLeuPheValPheIleSerIle    29

Qy  199 GGATCCACCATCAACTGGGT------GGAACAGAAAAGCCTTACCGGTCGACATG 249
        |||||| ::: :::  ||||       |||
Db   30 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAsnAsn 49

Qy  250 GTTCTCATCTCCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
        |||||| ||| ::: ||||       :::        |||
Db   50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis  69

Qy  310 ATCAGCGGTGGCCACCATCAACCCTGCCAGTGACTGGTGCCATGGTGCACCAGGAAGATC 369
        ||||||:: :::  ||||       :::  :::  ::||||
Db   70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle     89

Qy  370 AGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCGGGGGCCATCATTGGAGCA  429
        ||||| ::::::::  ||||| ||||||||||||||:::
Db   90 SerIlePheArgAlaLeuMetTyrIleIleAlaAlaGlnCysValGlyAlaIleValAlaThr 109

Qy  430 GGAATCCCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTGGGAGTCACCATGGTT 489
        |||||| :::    ||||:::  |||||||
Db  110 AlaIleLeuSerGlyIleThrSerSerLeuGlyGlyArgAsnSerLeuGlyArgAsnAspLeu 129

Qy  490 CATGGAAATCTTACCGGTGTCTCCTGGTGAGTTGATGATAATCACATTTCAATTG 549
        ||||| :::    ||||:::||||||
Db  130 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu  149

Qy  550 GTGTTTACTATCTTTGCCAGCTGTGATTCAAACGGACTGATGTCACTGGTCAATAGCT 609
        |||    ||| :::  ::: :::   :::||||
Db  150 ValLeuCysValLeuAlaThrThrAspArgArgArgArgAspLeuGlyGlySerAlaPro 169

Qy  610 TTAGCAATTGGATTTTCTGTTGCAATTGACATTTATTGCAATCAATTATACTGGTGCC 669
        |||||| :::    ||||:::||||
Db  170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuLeuAlaIleAspPheTyrThrGlyCys 189

Qy  670 AGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATCATGGAAAACCAT 729
        ::::||||||||||||:::     |||   :::  :::
Db  190 GlyIleGluAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209

Qy  730 TGGATATATGGGTTGGGCCCCATCATAGGAGGCTGTCCTCCGCTGGTGGCCTTTATGAGTAT 789
        ||||:::::||||:::||||       ||||||
Db  210 TrpIlePheTrpValGlyProPheIleGlyValAlaLeuAlaValLeuIleTyrAspPhe  229

Qy  790 GTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCTTCAGCAAAGCTGCC 849
        |||    :::        |||
Db  230 IleLeuLeuAlaPro------ArgSerAspLeuThrAspArgArgValLys          243

Qy  850 CAGCAAACAAAAGGAAGCTAC--------ATGGAGGTGGAGGACGAACAGGAGTCAG 897
        |||   |||:::||     :::||||
Db  244 ValTrpThrSerGlyGlnValGluTyrAspLeuGluAspAlaAlaAspIleAsnSerArg 263

Qy  898 GTAGAGACCGATGACCTGATTCTAAAACCT 927
        :::|||||||
Db  264 ValGlu------MetLysPro             268
```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5934
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-5934

Alignment Scores:
Pred. No.:
Score: 1.54e-51
Percent Similarity: 561.00   Length: 269
Best Local Similarity: 60.74%   Matches: 118
Query Match: 43.70%   Conservative: 46
DB: 25.72%   Mismatches: 88
                          4   Indels: 18
                                         Gaps: 4

US-10-723-180-1 (1-1152) x US-09-949-016-5934 (1-269)

```
Qy  139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCCTCCAGCCTG 198
        ||||||||::||||| ||||||||||||||||||||       ::::||||:::
Db   10 PheTrpArgAlaValValAlaGluPheLeuAlaThrLeuPheValPheIleSerIle    29

Qy  199 GGATCCACCATCAACTGGGT------GGAACAGAAAAGCCTTACCGGTCGACATG 249
        |||||| ::: :::  ||||       |||
Db   30 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAsnAsn 49

Qy  250 GTTCTCATCTCCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
        |||||| ||| ::: ||||       :::        |||
Db   50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis  69

Qy  310 ATCAGCGGTGGCCACCATCAACCCTGCCAGTGACTGGTGCCATGGTGCACCAGGAAGATC 369
        ||||||:: :::  ||||       :::  :::  ::||||
Db   70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle     89

Qy  370 AGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCGGGGGCCATCATTGGAGCA  429
        ||||| ::::::::  ||||| ||||||||||||||:::
Db   90 SerIlePheArgAlaLeuMetTyrIleIleAlaAlaGlnCysValGlyAlaIleValAlaThr 109

Qy  430 GGAATCCCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTGGGAGTCACCATGGTT 489
        |||||| :::    ||||:::  |||||||
Db  110 AlaIleLeuSerGlyIleThrSerSerLeuGlyGlyArgAsnSerLeuGlyArgAsnAspLeu 129

Qy  490 CATGGAAATCTTACCGGTGTCTCCTGGTGAGTTGATGATAATCACATTTCAATTG 549
        ||||| :::    ||||:::||||||
Db  130 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu  149

Qy  550 GTGTTTACTATCTTTGCCAGCTGTGATTCAAACGGACTGATGTCACTGGTCAATAGCT 609
        |||    ||| :::  ::: :::   :::||||
Db  150 ValLeuCysValLeuAlaThrThrAspArgArgArgArgAspLeuGlyGlySerAlaPro 169

Qy  610 TTAGCAATTGGATTTTCTGTTGCAATTGACATTTATTGCAATCAATTATACTGGTGCC 669
        |||||| :::    ||||:::||||
Db  170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuLeuAlaIleAspPheTyrThrGlyCys 189

Qy  670 AGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATCATGGAAAACCAT 729
        ::::||||||||||||:::     |||   :::  :::
Db  190 GlyIleGluAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209

Qy  730 TGGATATATGGGTTGGGCCCCATCATAGGAGGCTGTCCTCCGCTGGTGGCCTTTATGAGTAT 789
        ||||:::::||||:::||||       ||||||
Db  210 TrpIlePheTrpValGlyProPheIleGlyValAlaLeuAlaValLeuIleTyrAspPhe  229

Qy  790 GTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCTTCAGCAAAGCTGCC 849
        |||    :::        |||
Db  230 IleLeuLeuAlaPro------ArgSerAspLeuThrAspArgArgValLys          243

Qy  850 CAGCAAACAAAAGGAAGCTAC--------ATGGAGGTGGAGGACGAACAGGAGTCAG 897
        |||   :::|||     :::|||
```

RESULT 12
US-09-949-016-5934
; Sequence 5934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
Db    244 ValTrpThrSerGlyGlnValGluTyrAspLeuAspAlaAspIleAsnSerArg 263

Qy    898 GTAGAGACGGATGACCTGATTCTAAAACCT 927
          ||:||||||
Db    264 ValGlu----------MetLysPro 268

RESULT 13
US-09-949-016-8317
; Sequence 8317, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8317
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8317

- Alignment Scores:
  Pred. No.:                              Length:         281
  Score:                 1.57e-51         Matches:        118
  Percent Similarity:    561.00           Conservative:   46
  Best Local Similarity: 60.74%           Mismatches:     88
  Query Match:           43.70%           Indels:         18
  DB:                    25.72%           Gaps:           4
                         4

US-10-723-180-1 (1-1152) x US-09-949-016-8317 (1-281)

Qy    139 TTCTGGAAAGCAGTCACGCGGAATTTCTGGCCATGCTTATTTTGTTCCTCCAGCCTG 198
          ||||||::|||| |||||||||           ::||||:::::
Db     22 PheTrpArgAlaValValAlaGluPheLeuAlaThrLeuValPheIleSerIle 41

Qy    199 GGATCCACCATCAACTCGGGT--------GGAACAGAAAAGCCTTTACCGGTCGACATG 249
          ||||||  ::: ||||||         :::
Db     42 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 61

Qy    250 GTTCTCATCTCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
          ||| :::||| ||||||||| ||| |||:||
Db     62 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 81

Qy    310 ATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCACCAGGAAGATC 369
          |||||| :|:: ||:||| ::: :||:
Db     82 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle 101

Qy    370 AGCATCGCCAAGTGTCTTCTACATCGACCGAGCCCAGTGCCTCGGGGCCATCATTGGAGCA 429
          ||||   :::::||||:||:||       ::||||||
Db    102 SerIlePheArgAlaLeuMetTyrIleIleGluIleAlaGlnCysValGlyValAlaIleGlyAlaAlaThr 121

Qy    430 GGAATCCCTCTATCTGGTCACACCTCCCAGTGTGTGGGAGGCCTGGGAGTCACCATGGTT 489
          ||| |||||          ::: 
Db    122 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 141

Qy    490 CATGGAAATCTTACCGGTCGGTCATGGTCCCTGGTTGAGTTGATAATCACATTTCAATTG 549
          ||| ::: ||||:||:::|::||: |||||
Db    142 AlaAspGlyValAsnSerGlyLeuIleLeuGlyLeuIleIleGlyThrLeuGlnLeu 161

Qy    550 GTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGCTCAATAGCT 609
          |||   ::||||:::|   ::||||: 
Db    162 ValLeuCysValLeuAlaThrThrAspArgArgArgArgAspLeuGlyLeuGlySerAlaPro 181
```

```
Qy    610 TTAGCAATTGGATTTCGTTGCAATTGACATTTATTGCAATCAATTATACTGGTGCC 669
          ||||||||||| |||:|||||||||||:::|||||||||||
Db    182 LeuIleIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 201

Qy    670 AGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAATTGGGAAAACCAT 729
          :::||||||||||||||||||||||||||||||||||||:::||||
Db    202 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 221

Qy    730 TGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCCGCTGGTGGCCCTTTATGAGTAT 789
          ||||||||||||||||||| ||||||
Db    222 TrpIlePheTrpValGlyProPheIleGlyValGlyAlaLeuAlaLeuIleTyrAspPhe 241

Qy    790 GTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCCTTCAGCAAAGCTGCC 849
          :::                         ::: :::
Db    242 IleLeuAlaPro----------ArgSerAspLeuThrAspArgValLys 255

Qy    850 CAGCAAACAAAAGGAAGCTAC----------ATGGAGGTGAGGACAACAGGAGTCAG 897
          ||| |||          :::::: ::|||
Db    256 ValTrpThrSerGlyGlnValGluGlyTyrAspLeuAspAlaAspIleAsnSerArg 275

Qy    898 GTAGAGACGGATGACCTGATTCTAAAACCT 927
          ::|||||
Db    276 ValGlu----------MetLysPro 280

RESULT 14
US-09-949-016-10471
; Sequence 10471, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10471
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10471

Alignment Scores:
  Pred. No.:                              Length:         281
  Score:                 1.57e-51         Matches:        118
  Percent Similarity:    561.00           Conservative:   46
  Best Local Similarity: 60.74%           Mismatches:     88
  Query Match:           43.70%           Indels:         18
  DB:                    25.72%           Gaps:           4
                         4

US-10-723-180-1 (1-1152) x US-09-949-016-10471 (1-281)

Qy    139 TTCTGGAAAGCAGTCACGCGGAATTTCTGGCCATGCTTATTTTGTTCCTCCAGCCTG 198
          ||||||::|||| |||||||||           ::||||:::::
Db     22 PheTrpArgAlaValValAlaGluPheLeuAlaThrLeuValPheIleSerIle 41

Qy    199 GGATCCACCATCAACTCGGGT--------GGAACAGAAAAGCCTTTACCGGTCGACATG 249
          ||||||  ::: ||||||         :::
Db     42 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 61

Qy    250 GTTCTCATCTCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
          ||| :::||| ||||||||| ||| |||:||
Db     62 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 81

Qy    310 ATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGGCCATGGAAGATC 369
          |||||| :|:: ||:||| ::: :||:
Db     62 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 81
```

```
;  REGISTRATION NUMBER: 29,959
;  REFERENCE/DOCKET NUMBER: 5100-0003.00
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202) 887-1500
;     TELEFAX:  (202) 887-0763
;     TELEX:    90-4030
; INFORMATION FOR SEQ ID NO: 5:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 269 amino acids
;        TYPE: amino acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
US-08-447-554-5
```

```
Alignment Scores:
Pred. No.:               1.97e-51          Length:           269
Score:                   560.00            Matches:          118
Percent Similarity:      60.74%            Conservative:      46
Best Local Similarity:   43.70%            Mismatches:        88
Query Match:             25.68%            Indels:            18
DB:                                        Gaps:               4
```

US-10-723-180-1 (1-1152) x US-08-447-554-5 (1-269)

```
Qy    139 TTCTGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTTATTTTGTTCCTCAGCCTG 198
              ||| :: |||||  ||:|||          :::|||| :
Db     10 PheTrpArgAlaValAlaAlaThrThrLeuPheValPheLeuSerIle 29

Qy    199 GGAATCCACCATCAACTGGGT------------GGAACAGAAAAGCCTTACCGGTCGACATG 249
          ||||||||||          :::||||||
Db     30 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 49

Qy    250 GTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
          :: :|||||| ||||||||  :
Db     50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69

Qy    310 ATCAGGCGGTGGCCACATCAACCCTGCAGTGACTGGCCATGGTGTGCACCAGGAAGATC 369
          ||||||||| :::|| :|| :::|||
Db     70 IleSerGlyAlaHisLeuAsnProAlaValThrGlyHisGlyLeuLeuSerCysGlnIle 89

Qy    370 AGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGCCATCATTGGAGCA 429
          :||||||||| :   |||||||     ||||| |||:|||
Db     90 SerIlePheArgAlaLeuMetTyrIleIleAlaAlaGlnCysValGlyAlaIleIleAlaThr 109

Qy    430 GGAATCCTCTATCTGCTCACACTCCCCAGTGTGGTGGAGGCCTGGAGTCACCATGGTT 489
          ||||||||   :: :||
Db    110 AlaIleLeuSerGlyIleLeuThrSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 129

Qy    490 CATGGAAATCTTACCGGTCATGGTCTCCTGGTTCATGGTTGATAATCACATTCCAATTG 549
          ::: ||||| ||:: |:||||| ||||
Db    130 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleLeuGlyIleLeuIleGlyThrLeuGlnLeu 149

Qy    550 GTGTTTACTATCTTTGCCAGCTGTGATTCCAAAGGACTGATGTCACTGGCTCAATAGCT 609
          || |||||||||   :||:| ||||| :||
Db    150 ValLeuCysValLeuAlaThrThrArgArgArgAspLeuGlyLeuGlySerAlaPro 169

Qy    610 TTAGCAATTGGATTTCTGTTGCAATTGCAATTATTGCAATCAATTATACTGGTGCC 669
          ||||||||||||||||||   :: :||||| :| :||||||
Db    170 LeuAlaIleGlyLeuLeuSerValAlaLeuAspHisLeuLeuAlaIleLysAspTyrThrGlyCys 189

Qy    670 AGCATGAATCCCGCCGATCCTTTGACCTGCAGTTATCATGGGAAATTGGGAAAACCAT 729
          ::|||||||||||||||||||  ||||||||
Db    190 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209

Qy    730 TGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTGCTGGCCTTTATGAGTAT 789
          ||||||:|||||||||||||||   ||||||||
Db    210 TrpIlePheTrpValGlyProPheIleGlyValGlyProIleGlyAlaValLeuLeuAlaPheMetSerPhe 229

Qy    790 GTCTTCGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCCTTCAGCAAGCTGCC 849
          ::: ||| ||    :::|||||||
Db    230 IleLeuAlaPro----------ArgSerAspLeuThrAspArgValAsn 243

Qy    850 CAGCAAACAAAGGAAGCTAC--------------ATGGAGGTGGAGGACAACAGGAGTCAG 897
```

```
Db     82 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuLeuSerCysGlnIle 101

Qy    370 AGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGCCATCATTGGAGCA 429
              :: ||||||   :::::::          ||||||||
Db    102 SerIlePheArgAlaLeuMetTyrIleIleAlaAlaGlnCysValGlyAlaIleIleAlaThr 121

Qy    430 GGAATCCTCTATCTGCTCACACTCCCCAGTGTGGTGGAGGCCTGGAGTCACCATGGTT 489
          ||||||
Db    122 AlaIleLeuSerGlyIleLeuThrSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 141

Qy    490 CATGGAAATCTTACCGGTCATGGTCTCCTGGTTCATGGTTGATAATCACATTCCAATTG 549
              |||||||||||||||||||  :::
Db    142 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleLeuGlyIleLeuIleGlyThrLeuGlnLeu 161

Qy    550 GTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCACTGGCTCAATAGCT 609
          |||    ::: |||||:   ::|| :|||||    ||||
Db    162 ValLeuCysValLeuAlaThrThrArgArgArgAspLeuGlyLeuGlySerAlaPro 181

Qy    610 TTAGCAATTGGATTTCTGTTGCAATTGGACATTTATTGCAATCAATTATACTGGTGCC 669
          ||||||||||||||||||   :::||||| :||||||
Db    182 LeuAlaIleGlyLeuLeuSerValAlaLeuAspGluHisHisLeuLeuAlaIleLysAspTyrThrGlyCys 201

Qy    670 AGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAATGGGAAAACCAT 729
          ::: |||||||||||||||||||    ||| :||||||
Db    202 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 221

Qy    730 TGGATATATTGGGTTTGTGGGCCCATCATAGGAGCTGTCCTCTCCGTGGTGGCCTTTATGAGTAT 789
          ||||||||||||||||||   ||||||
Db    222 TrpIlePheTrpValGlyProPheIleGlyIleGlyGlyIleGlyAlaValLeuLeuAlaLeuValLeuIleTyrAspPhe 241

Qy    790 GTCTTCGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCCTTCAGCAAGCTGCC 849
          ::: ||| :|| :|||||||
Db    242 IleLeuAlaPro------------ArgSerAspLeuThrAspArgValLys 255

Qy    850 CAGCAAACAAAGGAAGCTAC--------------ATGGAGGTGGAGGACAACAGGAGTCAG 897
             ||| |||||||||
Db    256 ValTrpThrSerGlyGlnValGluGluTyrAspLeuAspAlaAlaAspAspIleAsnSerArg 275

Qy    898 GTAGAGACGGATGACCTGATTCTCAAAACCT 927
              :::|||||
Db    276 ----------MetLysPro 280
```

```
RESULT 15
US-08-447-554-5
; Sequence 5, Application US/08447554
; Patent No. 5661003
; GENERAL INFORMATION:
;  APPLICANT: FUSHIMI, KIYOHIDE
;  APPLICANT: UCHIDA, SHINICHI
;  APPLICANT: SASAKI, SEI
;  APPLICANT: MARUMO, FUMIAKI
;  TITLE OF INVENTION: WATER CHANNEL
;  NUMBER OF SEQUENCES: 8
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Morrison & Foerster
;     STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
;     CITY: Washington, DC
;     COUNTRY: USA
;     ZIP: 20006-1812
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/447,554
;     FILING DATE:
;     CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/126,365
;     FILING DATE: 24-SEP-1993
;  ATTORNEY/AGENT INFORMATION:
;     NAME: Murashige, Kate H.
```

```
Db     244 ValTrpThrSerGlyGlyValGluGluTyrAspLeuAspTyrAspAlaAspIleAsnSerArg 263
               |||   |||          ::::::    :::|||  |||::::
Qy     898 GTAGAGACGGATGACCTGATTCTAAAACCT 927
               ||||||          ::|||||||
Db     264 ValGlu--------------MetLysPro 268
```

Search completed: September 8, 2005, 12:16:57
Job time : 48 secs

OM nucleic - protein search, using frame_plus_n2p model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 8, 2005, 11:59:33 ; Search time 192 Seconds
(without alignments)
4725.879 Million cell updates/sec

Title: US-10-723-180-1
Perfect score: 2181
Sequence: 1 gggcaggcaatgagagctg.......attactcagtctaaacaata 1152

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop   6.0 , Delext   7.0

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 3548624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10723180/runat_08092005_123257_29276/app_query.fasta_1.1351
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10723180 @CGN_1_1.199 @runat_08092005_123257_29276
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1685	77.3	323	15	US-10-295-027-254	Sequence 254, App
2	1634	74.9	323	17	US-10-732-923-15771	Sequence 15771, A
3	1628	74.6	323	17	US-10-732-923-15770	Sequence 15770, A
4	1623	74.4	323	17	US-10-732-923-15775	Sequence 15775, A
5	1574	72.2	323	17	US-10-732-923-15576	Sequence 15576, A
6	1567	71.8	323	17	US-10-732-923-15577	Sequence 15577, A
7	1542.5	70.7	322	17	US-10-732-923-15579	Sequence 15579, A
8	1524	69.9	355	17	US-10-732-923-15574	Sequence 15574, A
9	1516	69.5	301	17	US-10-732-923-15772	Sequence 15772, A
10	1499.5	68.8	354	17	US-10-732-923-15578	Sequence 15578, A
11	1483	68.0	301	17	US-10-732-923-15575	Sequence 15575, A
12	1458.5	66.9	300	17	US-10-732-923-15580	Sequence 15580, A
13	656.5	30.1	249	17	US-10-732-923-15746	Sequence 15746, A
14	625	28.7	251	17	US-10-732-923-15615	Sequence 15615, A
15	617	28.3	265	14	US-10-097-340-12	Sequence 12, Appl
16	617	28.3	265	14	US-10-171-311-16	Sequence 16, Appl
17	617	28.3	265	17	US-10-645-756-6	Sequence 6, Appli
18	617	28.3	265	18	US-10-756-149-4696	Sequence 4696, Ap
19	606	27.8	263	17	US-10-732-923-15773	Sequence 15773, A
20	605	27.7	273	17	US-10-732-923-15761	Sequence 15761, A
21	604	27.7	271	17	US-10-732-923-15583	Sequence 15583, A
22	602	27.6	262	17	US-10-732-923-15760	Sequence 15760, A
23	602	27.6	263	17	US-10-732-923-15763	Sequence 15763, A
24	601	27.6	271	17	US-10-732-923-15584	Sequence 15584, A
25	599	27.5	262	17	US-10-732-923-15765	Sequence 15765, A
26	597.5	27.4	263	17	US-10-732-923-15581	Sequence 15581, A
27	595.5	27.3	263	17	US-10-732-923-15582	Sequence 15582, A
28	595	27.3	274	17	US-10-732-923-15762	Sequence 15762, A
29	589	27.0	271	18	US-10-794-899-95	Sequence 95, Appl
30	561	25.7	269	16	US-10-476-728A-6	Sequence 6, Appli
31	530.5	24.3	251	17	US-10-732-923-15661	Sequence 15661, A
32	525.5	24.1	222	17	US-10-645-756-10	Sequence 10, Appl
33	515	23.6	255	17	US-10-732-923-15667	Sequence 15667, A
34	508	23.3	282	15	US-10-403-161-10	Sequence 10, Appl
35	496.5	22.8	239	17	US-10-732-923-15747	Sequence 15747, A
36	496.5	22.8	245	17	US-10-732-923-15748	Sequence 15748, A
37	449.5	20.6	271	17	US-10-732-923-15745	Sequence 15745, A
38	447	20.5	72	9	US-09-864-761-3578	Sequence 3578, A
39	440	20.2	167	15	US-10-188-186-50	Sequence 50, Appl
40	438	20.1	248	16	US-10-437-963-178122	Sequence 178122,
41	437	20.0	290	16	US-10-437-963-156669	Sequence 156669,
42	436.5	20.0	210	16	US-10-645-756-8	Sequence 8, Appli
43	435	19.9	248	16	US-10-767-701-46122	Sequence 46122, A
44	434	19.9	247	15	US-10-424-599-191248	Sequence 191248,
45	434	19.9	247	17	US-10-732-923-15685	Sequence 15685, A

ALIGNMENTS

RESULT 1
US-10-295-027-254

; Sequence 254, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13

```
Db  161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180

Qy  580 AAACGGACTGATGTCACTGGCTCAATAAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA 639
        ||||||||
Db  181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200

Qy  640 CATTTATTGCAATCAATTATACTGGTGCCAGCATGAATCCGCCGATCCTTTGGACCT 699
        
Db  201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220

Qy  700 GCAGTTATCATGGGAAAATCCATTGGAAAACCATTGGATATATGGGTTGGGCCCATCATAGGA 759
        
Db  221 AlaValIleMetGlyAsnHisTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240

Qy  760 GCTGTCCTGCTGGTGGCCTTTATGAGATGTCTTCTGTCCAGATGTGAATTCAAACGT 819
        
Db  241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysValProAspValGluPheLysArg 260

Qy  820 CGTTTTAAAGAAGCTTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAGGTG 879
        
Db  261 ArgPheLysGluAlaPheSerLysAlaLeuAlaGlnInThrLysGlySerTyrMetGluVal 280

Qy  880 GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGAGTGGTGCAT 939
        
Db  281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300

Qy  940 GTGATTGACGTTGACCGGGGAGGAGGAGGAAGGGGAAAGACCAATCTCGAGAGGTATTG 999
        
Db  301 ValIleAspValAspArgGlyGluGluLysGlyLysAspGlnSerGlyGluValLeu 320

Qy 1000 TCTTCAGTA 1008
        |||||||
Db  321 SerSerVal 323
```

RESULT 2
```
US-10-732-923-15771
; Sequence 15771, Application US/10732923
; Publication No. US2005010879141
; GENERAL INFORMATION:
;   APPLICANT: Edgerton, Michael D
;   TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;   FILE REFERENCE: 38-15(52796)C
;   CURRENT APPLICATION NUMBER: US/10/732,923
;   CURRENT FILING DATE: 2003-12-10
;   PRIOR APPLICATION NUMBER: 10/310,154
;   PRIOR FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15771
;   LENGTH: 323
;   TYPE: PRT
;   ORGANISM: Bos taurus
US-10-732-923-15771
```

Alignment Scores:
```
Pred. No.:                           Length:        323
Score:                     1634.00   Matches:       312
Percent Similarity:        97.83%    Conservative:  4
Best Local Similarity:     96.59%    Mismatches:    7
Query Match:               74.92%    Indels:        0
DB:                                  Gaps:          0
```
1.61e-149

US-10-732-923-15771 (1-1152) x US-10-732-923-15771 (1-323)

```
Qy  40 ATGAGTGACGAGACCCACAGCAAGCCGGTGGGTAAGTGGGACCTTTGTGTTACCAGAGAG 99
        |||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetSerAspArgProAlaAlaArgProAlaAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20

Qy 100 AACATCATGTGGCTTTCAAAGGGGTCTGACTCAAGCTTCTGGAAAGCAGTCACAGCG 159
        ::::
Db  21 SerIleMetValIleAlaAlaPheLysGlyValTrpThrGlnThrPheThrLysAlaValThrAla 40

Qy 160 GAATTTCTGCCATGCTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGT 219
        
```

```
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 254
;   LENGTH: 323
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-254
```

Alignment Scores:
```
Pred. No.:                           Length:        323
Score:                     1685.00   Matches:       323
Percent Similarity:        100.00%   Conservative:  0
Best Local Similarity:     100.00%   Mismatches:    0
Query Match:               77.26%    Indels:        0
DB:                                  Gaps:          0
```
1.77e-154

US-10-723-180-1 (1-1152) x US-10-295-027-254 (1-323)

```
Qy  40 ATGAGTGACGAGACCCACAGCAAGCCGGTGGGTAAGTGGGACCTTTGTGTTCCAGAGAG 99
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetSerAspArgProThrAlaArgProThrAlaArgArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20

Qy 100 AACATCATGTGGCTTTCAAAGGGGTCTGACTCAAAGCTTCTCGGAAAGCAGTCACAGCG 159
        
Db  21 AsnIleMetValAlaAlaPheLysGlyValTrpThrGlnThrPheThrLysAlaValThrAla 40

Qy 160 GAATTTCTGGCCATGCTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGT 219
        
Db  41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60

Qy 220 GGAACAGAAAAGCCTTTACCGGTCGACATGGTTTCTCATCTCCCTTGCTTGGACTCAGC 279
        
Db  61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuCysLeuAspSerLeuSer 80

Qy 280 ATTGCAACCATGGTGCAGTGCTTTGGCCATCAGCGGTGGCCACATCAACCCTGCAGTG 339
        
Db  81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100

Qy 340 ACTGTGTGGCCATGTGTGCACCAGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA 399
        
Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleAla 120

Qy 400 GCCCAGTGCCTGGGGCCATCATTGGAGCGAGGAATCCTCTATCTGGTCACACCTCCCAGT 459
        
Db 121 AlaGlnCysLeuGlyLeuAlaIleIleGlyAlaGlyTyrLeuTyrLeuValThrProProSer 140

Qy 460 GTGGTGGGAGGGCCTGGAGTCACCATGTTCATTGGAAATCTTACCGCTGGTCATGGTCTC 519
        
Db 141 ValValGlyGlyLeuGlyLeuValThrMetValHisGlyValAsnLeuThrAlaGlyHisGlyLeu 160

Qy 520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC 579
```

```
Db   41  GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly   60

Qy  220  GGAACAGAAAAGCCTTACCGGTCGACATGGTTCCTCATCCCCTTGCTTTGGACTCAGC  279
         |||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GlyAlaGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer  80

Qy  280  ATTGCAACCATGGTGCAGTCGTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG  339
         ||||||||||||||||||||||||||||    |||||||||||||||||||||
Db   81  IleAlaThrMetValGlnSerPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal  100

Qy  340  ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA  399
         ||| |||||||||||||||||||||||||||   ::|||||||||||||||||
Db  101  ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleAla  120

Qy  400  GCCCAGTGCCTGGGGGCCCATCATTGGAGCAGGAGGAATCCTCATCTGGTCACCCCCAGT  459
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AlaGlnCysLeuGlyAlaHisHisTrpSerArgArgIleLeuIleTrpSerProProSer  140

Qy  460  GTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC  519
         ||| ||||||||||||||||||||||   :::|||||||||||||||||||||
Db  141  ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuSerAlaGlyHisGlyLeu  160

Qy  520  CTGGTTGAGTGATAATCACATTTCAATTGGTGTTTACTATCTTTGCCAGCTGTGATTCC  579
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy  580  AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyLeuSerValAlaIleGly  200

Qy  640  CATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCGCCCGATCCTTTGGACCT  699
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy  700  GCAGTTATCATGGGAAAACCATTGGAGATATATTGGGTTGGGCCCATCATAGGA  759
         |||||||||||||||||||||||||||||||||||||||||||||
Db  221  AlaValIleMetGlyAsnTrpProAspValGlyIleTyrTrpIleIleGly  240

Qy  760  GCTGTCCTCGGTGGCCCTTATGAGTATGTCTTCTGTCCAGAATGTTGAATTCAAACGT  819
         |||||||||||||||||||::::|||||||||||||||||
Db  241  AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820  CGTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGGAAGCTACATGGAGGTG  879
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  261  ArgPheLysGluLysProSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy  880  GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTCAAAACCTGAGTGGTGCAT  939
         ||||||||||||||||||||||||||||||||||::::|||
Db  281  GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGluLeuValHis  300

Qy 1000  TCTTCAGTA  1008
         |||||||||
Db  321  SerSerVal  323
```

RESULT 3
US-10-732-923-15770

; Sequence 15770, Application US/10732923
; Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: US/10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 15770
LENGTH: 323

; TYPE: PRT
; ORGANISM: Bos taurus
US-10-732-923-15770

Alignment Scores:
```
Pred. No.:                                Length:          323
Score:                  6.17e-149         Matches:         311
                        1628.00           Conservative:      4
Percent Similarity:     97.52%            Mismatches:        8
Best Local Similarity:  96.28%            Indels:            8
Query Match:            74.64%            Gaps:              0
DB:                     17
```

US-10-723-180-1 (1-1152) x US-10-732-923-15770 (1-323)

```
Qy   40  ATGAGTGACAGACCCACAGCAAGCGGTGGGTAAGTGTGGACCTTGTGTTACCAGAGAG   99
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MetSerAspArgProAlaAlaThrArgTrpGlyLysCysThrArgGlu   20

Qy  100  ACATCATGGTGCTTCAAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCG  159
         :::||||||||||||||||||||||||||||||||||||||||||||
Db   21  SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla  40

Qy  160  GAATTCTGGCCATGCTGTTATTTTGTCTCCTCAGCCTGGGATCCACCATCAACTGGGT  219
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   41  GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly   60

Qy  220  GGAACAGAAAAGCCTTACCGGTCGACATGGTTCCTCATCCCCTTGCTTTGGACTCAGC  279
         ||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GlyAlaGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer  80

Qy  280  ATTGCAACCATGGTGCAGTCGTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG  339
         ||||||||||||||||||||||||||||    |||||||||||||||||||||
Db   81  IleAlaThrMetValGlnSerPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal  100

Qy  340  ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA  399
         ||| |||||||||||||||||||||||||||   ::|||||||||||||||||
Db  101  ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleAla  120

Qy  400  GCCCAGTGCCTGGGGGCCCATCATTGGAGCAGGAGGAATCCTCATCTGGTCACCTCCCAGT  459
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AlaGlnCysLeuGlyAlaHisHisTrpSerArgArgIleLeuIleTrpSerProProSer  140

Qy  460  GTGGTGGGAGGCCTGGGAGTCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC  519
         ||| |||||||||||||||||||||||:::|||||||||||||||||||||
Db  141  ValValGlyGlyLeuGlyValValHisGlyAsnLeuSerAlaGlyHisGlyLeu  160

Qy  520  CTGGTTGAGTGATAATCACATTTCAATTGGTGTGTTTACTATCTTTGCCAGCTGTGATTCC  579
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy  580  AAACGGACTGATGTCACTGGCTGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyLeuSerValAlaIleGly  200

Qy  640  GCTGTCCTCGGTGGCCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAAACGT  819
         |||||||||||||||||||||::::||||||||||||||||||||||
Db  241  AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820  GTGGTGGGAGGCCTGGGAGTCATGGTCATGGTTCATGGAAATCTTAGTGGTGTTTACTAT  519
         |||||||||||||||||||||||||||   :::|||||||||||||||||||||
Db  141  ValValGlyGlyLeuGlyValValHisGlyAsnLeuSerAlaGlyHisGlyLeu  160

Qy  520  CTGGTTGAGTGATAATCACATTTCAATTGGTGTGTTTACTATCTTTGCCAGCTGTGATTCC  579
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy  580  AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyLeuSerValAlaIleGly  200

Qy  640  CATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCGCCCGATCCTTTGGACCT  699
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy  700  GCAGTTATCATGGGAAAATTGGGAAATTGGAGATATATTGGGTTGGGCCCATCATAGGA  759
         |||||||||||||||||||||||||||||||||||||||||||||
Db  221  AlaValIleMetGlyAsnTrpProAspValGlyIleTyrTrpIleIleGly  240

Qy  760  GCTGTCCTCGGTGGCCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAAACGT  819
         |||||||||||||||||||||::::||||||||||||||||||||||
Db  241  AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820  GCTGTCCTCGGTGGCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAAACGT  819
         |||||||||||||||||||||::::||||||||||||||||||||||
Db  241  AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820  CGTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGGAAGCTACATGGAGGTG  879
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  261  ArgPheLysGluLysProSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy  880  GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTCAAAACCTGAGTGGTGCAT  939
         ||||||||||||||||||||||||||||||||||::::|||
Db  281  GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGluLeuValHis  300
```

```
Qy   940  GTGATTGACGTTGACCGGGAGAGGAGAAGAAGGGGAAAGACCAATCTGGAGAGGTATTG  999
          ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  VallleAspIleAspArgArgGlyGluLysLysGlyLysAspProSerGlyGluValLeu  320

Qy  1000  TCTTCAGTA  1008
          |||||||||
Db   321  SerSerVal  323
```

RESULT 4
US-10-732-923-15775
; Sequence 15775, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;   APPLICANT: Edgerton, Michael D
;   TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;   FILE REFERENCE: 38-15(52796)C
;   CURRENT APPLICATION NUMBER: US/10/732,923
;   CURRENT FILING DATE: 2003-12-10
;   PRIOR APPLICATION NUMBER: 10/310,154
;   PRIOR FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15775
;   LENGTH: 323
;   TYPE: PRT
;   ORGANISM: Ovis aries
US-10-732-923-15775

Alignment Scores:
Pred. No.:   Length: 323
Score: 1623.00   Matches: 310
Percent Similarity: 97.52%   Conservative: 5
Best Local Similarity: 95.98%   Mismatches: 8
Query Match: 74.42%   Indels: 0
DB: 17   Gaps: 0

US-10-723-180-1 (1-1152) x US-10-732-923-15775 (1-323)

```
Qy    40  ATGAGTGACAGACCCACAGCAGCAAGGCGGTGGGTAAGTGTGGGACCTTTGTGTACCAGAGAG   99
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MetSerAspArgProAlaAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu   20

Qy   100  AACATCATGTGGCTTCAAAGGGGTCTCGACTCAAGCTTCTCGAAAGCAGTCACAGCG  159
          :::|||||||||||||||||||||||||||||||||||||||||||:::|||||
Db    21  SerIleMetValAlaPheLysGlyValThrPheGlnThrPheLysAlaValAlaThrAla   40

Qy   160  GAATTCTGGCCATGCTTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGT  219
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41  GluPheLeuLeuAlaMetLeuLlePheValLeuLeuSerLeuSerThrIleAsnTrpGly   60

Qy   220  GGAACAGAAAAGCCTTACCGGTGACATGGTTCCATCTCCCTTGCTTTGGACTCAGC  279
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GlyAlaGluLysProLeuProValAspMetValLeuIleSerLeuCysPheLeuSer   80

Qy   280  ATTGCAACCATGGTGCAGTGCTTTGGCCACCATATACGGTGGCCACACCACCCGCAGTG  339
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81  IleAlaThrMetValGlnCysPheGlyHisHisIleArgTrpProHisHisProAlaVal  100

Qy   340  ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA  399
          |||||||||||||||||::|||||||||||||||||||||||||::|||||||||||||
Db   101  ThrValAlaMetValCysThrArgArgIleSerIleAlaLysSerValPheTyrIleAla  120

Qy   400  GCCCAGTGCCGTGGGGCCATCATTGGAGCAGGAATCCTCATCTGGTCACACCTCCCAGT  459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuIleTrpSerHisLeuProSer  140

Qy   460  GTGGTGGGAGGCCTTGGAGTCACCGTCACCGTCTTCATCCCGTGTCATGGTCTCC  519
          ||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db   141  ValValGlyGlyLeuGlyValThrValThrValHisArgAsnLeuSerAlaGlyHisGlyLeu  160

Qy   520  CTGGTTGAGTTGGATAATCACATTTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC  579
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161  LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180
```

```
Qy   580  AAACGGACTGATGTTCACTGCTCAATAGCTTTAGCAATTGGATTTCTGTTGCAATTGGA  639
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly  200

Qy   640  CATTTATTGCAATCAATTATACTGGTGCCAGCATGATGCCCGATCCTTGGACCT  699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy   700  GCAGTTATCATGGGAAATTGGGAAAACCATTGGATATATTGGGTTGGGCCATCATAGGA  759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221  AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly  240

Qy   760  GCTGTCCTGCTGTGGCCTTTATGAGTATGTCTTCGTCCAGATGTGAATTCAAACGT  819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  AlaValLeuAlaGlyAlaGlyLeuTyrLeuTyrGluThrValPheCysProAspValLeuLysArg  260

Qy   820  CGTTTAAAGAAGCCTTCAGCAAAGTCGCCCAGCAAACAAAAGGAAGCTACATGGAGGTG  879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   261  ArgPheLysGlyAlaAlaPheSerLysAlaAlaGlnGlnThrLysGlyGlySerTyrMetGluVal  280

Qy   880  GAGGACAACAGGAGTCAGGTAGGACAGCGGATGACCTGATTCTAAAACCTGGAGTGGTGCAT  939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   281  GluAspAsnArgSerGlnValGluThrArgAspAspLeuIleLeuLysProGlyValValHis  300

Qy   940  GTGATTGACGTTGACCGGGAGAGGAGAAGAAGGGGAAAGACCAATCTGGAGAGGTATTG  999
          ||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db   301  VallleAspIleAspArgArgGlyGluLysLysGlyLysAspProSerGlyGluValLeu  320

Qy  1000  TCTTCAGTA  1008
          |||||||||
Db   321  SerSerVal  323
```

RESULT 5
US-10-732-923-15776
; Sequence 15776, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;   APPLICANT: Edgerton, Michael D
;   TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;   FILE REFERENCE: 38-15(52796)C
;   CURRENT APPLICATION NUMBER: US/10/732,923
;   CURRENT FILING DATE: 2003-12-10
;   PRIOR APPLICATION NUMBER: 10/310,154
;   PRIOR FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15776
;   LENGTH: 323
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-732-923-15776

Alignment Scores:
Pred. No.:   Length: 323
Score: 1574.00   Matches: 301
Percent Similarity: 95.67%   Conservative: 8
Best Local Similarity: 93.19%   Mismatches: 14
Query Match: 72.17%   Indels: 0
DB: 17   Gaps: 0

US-10-723-180-1 (1-1152) x US-10-732-923-15776 (1-323)

```
Qy    40  ATGAGTGACAGACCCACAGCAGCAAGGCGGTGGGTAAGTGTGGGACCTTTGTGTACCAGAG   99
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MetSerAspArgAlaAlaAlaArgArgTrpGlyLysCysGlyHisSerArgGlu   20

Qy   100  AACATCATGTGGCTTCAAAGGGGTCTCGACTCAAGCTTCTCGAAAGCAGTCACAGCG  159
          :::|||||||||||||||||||||||||||||||||||||||||:::||||||
Db    21  SerIleMetValAlaPheLysGlyValThrPheGlnThrPheLysProLysAlaValSerAla   40

Qy   160  GAATTCTGGCCATGCTTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGT  219
          ||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db    41  GluPheLeuLeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGly   60
```

```
;  ORGANISM: Mus musculus
US-10-732-923-15577

Alignment Scores:
Pred. No.:
Score:                    5.26e-143       Length:          323
                          1567.00         Matches:         300
Percent Similarity:       95.36%          Conservative:    8
Best Local Similarity:    92.88%          Mismatches:      15
Query Match:              71.85%          Indels:          0
DB:                       17              Gaps:            0

US-10-723-180-1 (1-1152) x US-10-732-923-15577 (1-323)

Qy   40  ATGAGTGACAGACCACCAGCAAGGCGGTGGGTAAGTGGACCTTGTGTACCAGAGAG   99
         ||||||||
Db    1  MetSerAspGlyAlaAlaAlaAlaArgArgTrpGlyLysCysGlyHisSerCysSerArgGlu   20

Qy  100  AACACATCAGTGGCTTCAAAGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACACGCG  159
         :::||||||||||||
Db   21  SerIleMetValAlaPheLysGlyValTrpPheThrGlnAlaPheThrLysAlaValSerAla   40

Qy  160  GAATTCTGGCCATGGCTTATTTTGTTCTCCTCCAGCTGGATCCACCATCAACTGGGGT  219
         ||||||||
Db   41  GluPheLeuAlaThrLeuIlePheValLeuLeuLeuGlyValLeuGlySerThrIleAsnTrpGly   60

Qy  220  GGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTGGACTCAGC  279
         |||:::||| ||||||| ||||||||||||||||||
Db   61  GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer   80

Qy  280  ATTGCAACCATGGTCGAGTGCTTTGGCCATATCAGCGGTGCCACATCAACCCTGCAGTG  339
         |||||||| ||| ||||||||||||||||
Db   81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisGlyHisGlyVal  100

Qy  340  ACTGTGGCCATGGTGTGCACCAGAGAGATCAGCATCGCCCAAGTCTGTCTTCACATCGCA  399
         |||||||
Db  101  ThrValAlaMetValCysThrArgGluIleSerIleAlaLysSerIleSerValPheTyrIleIle  120

Qy  400  GCCCAGTGCCTGGGGCCCATCATTGGATATATGTTGGGCCCCATCATAGGA  459
         |||||||
Db  121  AlaGlnCysLeuGlyIleAlaIleIleGlyTyrLeuValThrProProSer  140

Qy  460  GTGGTTGGGGAGGCCTGGAGTCCACCATGGTTCATGGAAATCTTCATGGGTGTTGTTACATCTTCAC  519
         |||||||
Db  141  ValValGlyGlyLeuGlyValThrThrValHisGlyHisAlaGlyHisGlyLeu  160

Qy  520  CTGGTTGAGTTGATAATCACATTCAATTGGTGTGTTCTGGTTVHisIleAlaLeuHis  579
         |||||||
Db  161  LeuValGluLeuIleIleIleThrPheGlnLeuLeuPheThrIleuAlaSerCysAspSer  180

Qy  580  AAACCGACTGATGTCAATCGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
         |||||||
Db  181  LysArgThrAspValThrArgLeuAsnSerIleAlaLeuAlaIleGlyPheSerValAlaIleGly  200

Qy  640  CATTTATTTGCAATCAATTAATACTGGTGCCAGCATCATGAATCCCGCCGATCCTTTGGACCT  699
         |||||||
Db  201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy  700  GCAGTTATCATGGGAAATTGGGAAAACCATTGGAATATATGGGTTGGCCCATCATAGGA  759
         |||||||
Db  221  AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGly  240

Qy  760  GCTGTTCCTCGCTGGTTGGCCTGGGCCTTTATGAGTATGTCTTCTCGTCGTGGTCATCATCGTGGTCATGGTGCTCTC  819
         |||||||
Db  241  AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820  CGTTTTAAAGAAGCCTTGACCGGAGAGAAGAAGAAAGGGGAAAGACCAATCTGAGGAGTATTG  879
         |||||||
Db  261  ArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy  880  GAGGACAACAGGAGTCAGGTGAGACGGATGAGTCAATTGCTCGTTGCAATTGGA  939
         |||||||
Db  281  GluAspAsnArgSerGlnValGluThrGluAspGluLysLysGlyGluLeuVal  300

Qy  940  GTGATTGACGTTGACCCGGGAGGAGGAAGAAGGGGAAAGACCAATCTGAGGAGTATTG  999
         |||::||| :::||||||| ||||||||||||||
Db  301  ValIleAspValAspProGlyGlyGlyArgArgGlyGluGluLysGlyLysLysAspSerSerGlyGlyLeu  320

Qy 1000  TCTTCAGTA 1008
         |||||||
Db  321  SerSerVal 323
```

```
Qy  220  GGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTGGACTCAGC  279
         |||:::||| ||||||| ||||||||||||||||||
Db   61  GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer   80

Qy  280  ATTGCAACCATGGTCGAGTGCTTTGGCCATATCAGCGGTGCCACATCAACCCTGCAGTG  339
         |||||||| ||| ||||||||||||||||
Db   81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisGlyHisGlyVal  100

Qy  340  ACTGTGGCCATGGTGTGCACCAGAGAGATCAGCATCGCCCAAGTCTGTCTTCACATCGCA  399
         |||||||
Db  101  ThrValAlaMetValCysThrArgGluIleSerIleAlaLysSerIleSerValPheTyrIleIle  120

Qy  400  GCCCAGTGCCTGGGGCCCATCATTGGAGCAGAATCCTCATCTGTCACACCTCCCAGT  459
         |||||||
Db  121  AlaGlnCysLeuGlyIleAlaIleIleGlyTyrLeuValThrProProSer  140

Qy  460  GTGGTTGGGGAGGCCTGGAGTCCACCATGGTTCATGGAAATCTTACCGCTGGTCATGTGTTC  519
         |||||||
Db  141  ValValGlyGlyLeuGlyValThrThrValHisGlyHisAlaGlyHisGlyLeu  160

Qy  520  CTGGTTGAGTTGATAATCACATTTACTATCTTTGCCAGCTGTGATTCC  579
         |||||||
Db  161  LeuValGluLeuIleIleIleThrPheGlnLeuLeuPheThrIleuAlaSerCysAspSer  180

Qy  580  AAACCGACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
         |||||||
Db  181  LysArgThrAspValThrArgLeuAsnSerIleAlaLeuAlaIleGlyPheSerValAlaIleGly  200

Qy  640  CATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCT  699
         |||||||
Db  201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy  700  GCAGTTATCATGGGAAATTGGGCCTTTATGAGTATGTCTTCTCGTCGTGGTCATCATAGGA  759
         |||||||
Db  221  AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGly  240

Qy  760  GCTGTCCTCGCTGGTGGCCTGGGCCTTTATGAGTATGTCTTCTCGTCCAGATGTTGAATTCAAACGT  819
         |||||||
Db  241  AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820  CGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGGGAAGCTACATGGAGGTG  879
         ||| ||||||| :::||||||| ||||||||||||||
Db  261  ArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy  880  GAGGACAACAGGAGTCAGGTGAGACGGATGAGTCAATTGCTCGTTGCAATTGGA  939
         |||||||
Db  281  GluAspAsnArgSerGlnValGluThrGluAspGluLysLysGlyGluLeuVal  300

Qy  940  GTGATTGACGTTGACCCGGGAGGAGGAAGAAGGGGAAAGACCAATCTGAGGAGTATTG  999
         |||::||| :::||||||| ||||||||||||||
Db  301  ValIleAspValAspArgArgGlyGluGluLysGlyLysLysAspSerSerGlyGlyLeu  320

Qy 1000  TCTTCAGTA 1008
         |||||||
Db  321  SerSerVal 323
```

RESULT 6
US-10-732-923-15577
; Sequence 15577, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15577
; LENGTH: 323
; TYPE: PRT

```
Qy   580 AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA 639
             ||||||||||||||||||||::||||||||||||  ||||||||||||||||||||
Db   180 LysArgThrAspValThrGlyGlySerIleAlaLeuAlaIleAlaIleGlyPheSerValAlaIleGly 199

Qy   640 CATTTATTTGCAATCAATTATACTGGTGCCAGCAGATGAATCCGCCGATCCTTGGACCT 699
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   200 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 219

Qy   700 GCAGTTATCATGGGAAATTGGGAAAACCATTGGATAATATTGGGTTGGGCCCATCATAGGA 759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGly 239

Qy   760 GCTGTCCTGCTGGCCCTTATGAGTAGTGTCTTCTGTCCAGATGTTGAATTCAAACGT 819
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLysArg 259

Qy   820 CGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAGGTG 879
             |||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   260 ArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 279

Qy   880 GAGGACAACAGGAGTCAGGTAGAGACGGATGAGTCTAAAACCTGAGTGGTGCAT 939
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280 GluAspAsnArgSerGlnValGluThrAspGluSerLysThrGluLysProGlyValValHis 299

Qy   940 GTGATTGACGTTGACCGGGAGAGAGAAGGGGAAAGACCAATCTGGAGAGGTATTG 999
             ||||||||||||||||||||||||||||||||||||||||||||||||||:::||
Db   300 ValIleAspValAspArgGlyGluLysLysGlyLysAspSerGlyGluValLeu 319

Qy  1000 TCTTCAGTA 1008
             |||||||||
Db   320 SerSerVal 322
```

RESULT 8
US-10-732-923-15574
; Sequence 15574, Application US/10732923
; Publication No. US2005010879A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15574
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-732-923-15574

Alignment Scores:

Pred. No.:		Length:	355
Score:	8.24e-139	Matches:	292
Percent Similarity:	1524.00	Conservative:	8
Best Local Similarity:	96.15%	Mismatches:	12
Query Match:	93.59%	Indels:	0
DB:	69.88%	Gaps:	0
	17		

US-10-723-180-1 (1-1152) x US-10-732-923-15574 (1-355)

```
Qy    73 AAGTGTGGACCTTTGTACCAGAGAGAACATCATGGTGGCTTTCAAAGGGGTCGGGACT 132
             ||||:|||||||  |||:::|||||||||||||||||||||||||||||||||||
Db    44 LysCysGlyHisSerCysSerArgGluSerIleMetValAlaPheLysGlyValTrpThr 63

Qy   133 CAAGCTTTCTGGAAAGCAGTCACACGGAATTTCGGCCATGCTATTTTGTTCTCCTC 192
             ||:::||||||||||||||||||||||||||||||||||||||  ||||||||||
Db    64 GlnAlaPheTrpLysAlaValSerAlaGluPheLeuAlaThrLeuIlePheValLeu 83

Qy   193 AGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTT 252
             ||::||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db    84 GlyValGlySerThrIleAsnTrpGlySerGluAsnProLeuProValAspAspMetVal 103
```

```
Score:                 1516.00              Matches:        292
Percent Similarity:    98.01%               Conservative:     3
Best Local Similarity: 97.01%               Mismatches:       6
Query Match:           69.51%               Indels:           0
DB:                                         Gaps:             0

US-10-723-180-1 (1-1152)  x  US-10-732-923-15772 (1-301)

Qy    106 ATGGTGGCTTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTT 165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAlaGluPhe   20

Qy    166 CTGGCCATGCTTATTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACA 225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 LeuAlaMetLeuIlePheValLeuLeuSerSerAlaTrpAspProProSerThrGlyGlyThrIleAsnTrpGlyGlyAla   40

Qy    226 GAAAGCCTTACCGGTCGTCATGGTTGATAATCACATTCAATTGGTG 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 GluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla   60

Qy    286 ACCATGGTGCAGTGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCCAGTGACTGTG 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrVal   80

Qy    346 GCCATGGTGTGCACCGGAGAGATCAGCATCCGCCAAGTCTGTTCTTCACATCGGTGCCAGC 405
          ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db     81 AlaMetValCysThrArgArgIleSerIleAlaLysSerValPheTyrIleAlaAlaGln  100

Qy    406 TGCCTGGGGCCATCATTGGACCTGCAGTTATCATGGGAAAATGGAAAACCATTGG 465
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 CysLeuGlyAlaIleIleGlyAlaGlyIleIleGluTyrLeuValThrProProSerValVal  120

Qy    466 GGAGGCCTGGGAGTGCCACCATGGTTCATGGAAAATCTTACCCGTCGGTCATGGTCTCCTGGTT 525
          ||||||||||||||||||||::||||||||||||||||||||:::||||||||||||||||
Db    121 GlyGlyLeuGlyValThrThrValHisGlyAsnLeuSerAlaGlyHisGlyLeuLeuVal  140

Qy    526 GAGTTGATAATCACATTCAATTGGTGTTTACTATCTTGCCAGCTGTGATTTTCGTGTGCAATTGGCAATTTA 585
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArg  160

Qy    586 ACTGATGTCACTGGCTCAATAGCTTTAGCAATTGATTTTCGTGTCAATTGGCAATTA 645
          ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db    161 ThrAspValThrGlySerIleAlaLeuAlaIleAspPheArgValAsnGlyHisLeu  180

Qy    646 TTTGCAATCAATTATACTGTGCCAGCATGCAATCCGCCCGATCCTTTGACCTGCAGTT 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaVal  200

Qy    706 ATCATGGGAAATTGGGAAAACCATTGGATATATTGGTTTGGGCCCATCATAGGAGCTGTC 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 IleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaVal  220

Qy    766 CTCGCTGGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTT 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 LeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArgArgPhe  240

Qy    826 AAAGAAGCCTTCAGCAAAGCTGCCAGCAAAGCTACCATGGAGGTGGAGGAC 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LysGluAlaPheSerLysAlaAlaSerLysAlaThrMetGluValGluAsp  260

Qy    886 AACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACTGGAGTGGTGCATGTGATT 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 AsnArgSerGlnValGluThrAspAspLeuIleLeuLysLeuGluTrpValHisValIle  280

Qy    946 GACGTTGACCGGGAGGAGAAGAAGGGAAAGACCAATCTGAGAGGGTATTGTCTTCA 1005
          ||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 AspIleAspArgArgGlyGluGluLysLysGlyLysLysAspProSerGlyGluValLeuSerSer  300

Qy   1006 GTA 1008
          |||
Db    301 Val 301
```

```
Qy    253 CTCATCTCCCTTTGCTTTGGACTCAAGCCATTGCAACCATGGTGCAGTGCTTTGGCCATATC 312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 LeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetValGlnCysPheGlyHisIle  123

Qy    313 AGCGGTGGCCACATCAACCCTGCCAGTGACTGTGGCCATGGTGTGCACCGGAAGATCAGC 372
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 SerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThrArgLysIleSer  143

Qy    373 ATCGGCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCCATCATTGGAGCAGGA 432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 IleAlaLysSerValPheTyrIleAlaAlaGlnCysLeuGlyAlaIleIleGlyAlaGly  163

Qy    433 ATCCTCTATCTGGTCACACTCCCAGTGGTGTGGGAGGCCTGGGAGTCACCATGGTTCAT 492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    164 IleLeuTyrLeuValThrProProSerValGlyGlyLeuGlyValThrThrValHis  183

Qy    493 GGAAATCTTACCGGTGGTCATGGTCCTGGTTGAGTTGATAATCACATTCAATTGGTG 552
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 GlyAsnLeuThrAlaGlyHisGlyLeuLeuValGluLeuIleIleThrPheGlnLeuVal  203

Qy    553 TTTACTATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGCTCAATAGCTTTA 612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 PheThrIlePheAlaSerCysAspSerLysArgLeuMetSerLeuAlaGlnSerIleAlaLeu  223

Qy    613 GCAATTGGATTTTCTGTTGCAATTGACATTTATTTGCAATCAATTATACTGTGCCAGC 672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    224 AlaIleGlyPheSerValAlaIleGlyIleMetGlyAsnLeuSerAlaGlyHisGlyLeu  243

Qy    673 ATGGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAATTGGAAAACCATTGG 732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 MetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTrpAlaAsnHisTrp  263

Qy    733 ATATATTGGGTTGGGCCCATCATAGGAGCTGTCTCCTCGCTGGTGGCCTTTATGAGTATGTC 792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264 IleTyrTrpValGlyProIleMetGlyIleAlaLeuAlaLeuAlaLeuTyrGluTyrVal  283

Qy    793 TTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGCCTTCAGCAAAGCCTGCCCAG 852
          ||||||||||||||||||||    |||||||||||||||||||||||||||||
Db    284 PheCysProAspValGluLeuLysArgArgLeuLysLysGluGlyLysLeuGluLysSerLysAlaAlaGln  303

Qy    853 CAAACAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGTCAGGTAGAGACGGATGAC 912
          ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||
Db    304 GlnThrGlySerTyrMetGluValGluAspAsnArgSerGlnValGluThrGluAsp  323

Qy    913 CTGATTCTAAAACCTGGAGTGGTGCATGTGATTGACGTTGACCGGGAGGAGAAGAAG 972
          |||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db    324 LeuIleLeuLysLysProGlyValValHisValIleAspIleAspArgGlyGluGluLysLys  343

Qy    973 GGGAAAGACCAATCTGAGAGGGTATTGTCTTCAGTA 1008
          ||||||||||||||||||||||||||||||||||||
Db    344 GlyLysAspSerGlySerGlyGluValLeuSerSerVal  355
```

RESULT 9
US-10-732-923-15772
; Sequence 15772, Application US/10732923
; Publication No. US20050108791A1
GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15772
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-732-923-15772

Alignment Scores:
Pred. No.:   4.65e-138   Length: 301

RESULT 10

US-10-732-923-15578
; Sequence 15578, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;   APPLICANT: Edgerton, Michael D
;   TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;   FILE REFERENCE: 38-15(52796)C
;   CURRENT APPLICATION NUMBER: US/10/732,923
;   CURRENT FILING DATE: 2003-12-10
;   PRIOR APPLICATION NUMBER: 10/310,154
;   PRIOR FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 24149
;   SEQ ID NO 15578
;   LENGTH: 354
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-732-923-15578

Alignment Scores:
Pred. No.:   Length: 354
Score: 1.98e-136   Matches: 288
Percent Similarity: 95.51%   Conservative: 10
Best Local Similarity: 92.31%   Mismatches: 13
Query Match: 68.75%   Indels: 1
DB: Gaps: 1

US-10-723-180-1 (1-1152) x US-10-732-923-15578 (1-354)

Qy   73 AAGTGTGGACCTTTGTGTACCAGAGAGAACATCATGGTGGCTTTCAAAGGGGTCTGGACT 132
                  ||||||| ||:::|||||||||||||||||||||||||||||||||||||||
Db   44 LysCysGlyHisSerCysSerArgGluSerIleMetValAlaPheLysGlyValTrpThr   63

Qy   133 CAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCTCCTC 192
                  |||||||||||||||| |||:::|||||||||||||||||||||||::|||||
Db   64 GlnAlaPheTrpLysAlaValSerAlaGluPheLeuAlaMetLeuIlePheVal--Leu   82

Qy   193 AGCCTGGATCCACCATCAACCTGGGGTGGAAACAGAAAAGCCTTTACCGGTCGACATGGTT 252
                  |||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db   83 GlyValGlySerThrIleAsnTrpGlyTrpLysGlnLysSerLeuTyrArgSerThrTrp 102

Qy   253 CTCATCTCCCTTTGCTTTGGACTTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCATATC 312
                  |||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db   103 LeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetValGlnCysLeuGlyHisIle 122

Qy   313 AGCGGTGGCCACATCAACCCTGCAGTGACTGCCATGGTGTGCACCAGGAAGATCAGC   372
                  |||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db   123 SerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThrArgLysIleSer 142

Qy   373 ATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGCCCATCATTGGAGCAGGA 432
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 IleAlaLysSerValPheTyrIleAlaAlaGlnCysLeuGlyAlaIleIleGlyAlaGly 162

Qy   433 ATCCTCTATCTGGTCACACCTCCCAGTGGTGGGAGGCCTGGGAGTCCACCATGGTTCAT 492
                  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   163 IleLeuTyrLeuValThrProSerValGlyGlyArgLeuGlyValValHisThrValHis 182

Qy   493 GGAAATCTACCGCTGGTCATGTCTCCTGGTGAGTGATAATCACATTCAATTGGTG 552
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 GlyAsnLeuThrAlaGlyHisGlyLeuValGluLeuIleThrPheGlnLeuVal 202

Qy   553 TTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCACTGGCTCAAATAGCTTTA 612
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 PheThrValPheAlaSerCysAspSerLysArgThrAspValThrGlySerIleAlaLeu 222

Qy   613 GCAATTGGATTTTCTGTTGCAATTGGACATTTATTTGCAATCAATTATACTGGTGCCAGC 672
                  ||||||||||||||||||||||||||||||||||||||||||||::::::|||||
Db   223 AlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTyrThrGlyAlaSer 242

Qy   673 ATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAATTGGGAAAACCATTGG 732
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 MetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTrpAlaAsnHisTrp 262

Qy   733 ATATATTGGGTTGGCCCATCATAGAGAGCTGTCCTCGCTGGCTGGTGCCTTTATGAGTATGTC 792

Db   263 IleTyrTrpValGlyProIleMetGlyAlaValLeuAlaGlyAlaLeuTyrGluTyrVal 282

Qy   793 TTCTGTCCAGATGTGAATTCAAACGTCGTTTTAAAGAAGAGCCTTCAGCAAAGCGCCCAG 852

Db   283 PheCysProAspValGluLeuLysArgArgLeuLysArgLeuLysAlaPheSerLysAlaAlaGln 302

Qy   853 CAAACAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGTCAGGTAGAGACGGATGAC 912

Db   303 GlnThrLysGlySerTyrMetGluValGluAspAsnArgSerGlnValGlyThrGluAsp 322

Qy   913 CTGATTCTAAAACCTGGAGTGGTGCATGTGATTGACGTTGACCGGGAGGAGAAGAAG 972

Db   323 LeuIleLeuLysProGlyValValHisValIleAspValAspArgGluGluLysLys 342

Qy   973 GGGAAAAGACCAATCTGGAGAGGGATTATTGTCTTCAGTA 1008
                  |||:::|||||||||||||||||||||||||||||
Db   343 GlyArgAspSerGlyGluValLeuSerSerVal 354

RESULT 11
US-10-732-923-15575
; Sequence 15575, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;   APPLICANT: Edgerton, Michael D
;   TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;   FILE REFERENCE: 38-15(52796)C
;   CURRENT APPLICATION NUMBER: US/10/732,923
;   CURRENT FILING DATE: 2003-12-10
;   PRIOR APPLICATION NUMBER: 10/310,154
;   PRIOR FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 24149
;   SEQ ID NO 15575
;   LENGTH: 301
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-732-923-15575

Alignment Scores:
Pred. No.:   Length: 301
Score: 7.51e-135   Matches: 285
Percent Similarity: 96.68%   Conservative: 6
Best Local Similarity: 94.68%   Mismatches: 10
Query Match: 68.00%   Indels: 0
DB: Gaps: 0

US-10-723-180-1 (1-1152) x US-10-732-923-15575 (1-301)

Qy   106 ATGGTGGCTTTCAAAGGGTCTCGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTT 165

Db   1 MetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValSerAlaGluPhe   20

Qy   166 CTGGCCATGCTTATTTTGTTCTCCCTCAGCCTGGATCCACCATCAACTGGGGTGGAAACA 225

Db   21 LeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGlyTrpGlySer   40

Qy   226 GAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTGCA 285

Db   41 GluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla   60

Qy   286 ACCATGGTGCAGTGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTG 345

Db   61 ThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrVal   80

Qy   346 GCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGC 405

Db   81 AlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleAlaAlaGln 100

Qy   406 TGCCTGGGGCCCATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCCAGTGGTGTG 465

Db   101 CysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSerGlyVal 120

Qy   466 GGAGGCCTGGGAGTCCATCATGGTTCATGGAAATCTTACCGCTGGTGCCTTTATGAGTATGTC 525

```
Db         ||||| :||||||||||||     |||      :::|||||||||||||||||||||:::
       21  LeuAlaThrLeuIlePheVal---LeuGlyValGlySerThrIleAsnTrpGlyGlySer  39

Qy    226  GAAAAGCCTTACCGGTCGACATGGTTCTCATCTCCTTTGACTGTGATTCAGCAATTGCA  285
           |||        ||||||||||||||||||||||||||||||||||||||||||||||
Db     40  GluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla  59

Qy    286  ACCATGGTGCAGTGCTTTGCCAATCAGCGGTGCCACATCAACCCTGCAGTGACTGTG  345
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60  ThrMetValGlnCysLeuGlyHisIleSerGlyHisIleAsnProAlaValThrVal  79

Qy    346  GCCATGGTGTGCACCAGGAGAGATCAGCATGCGCAAGTCTGTCTTCTACATCGCAGCCCAG  405
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80  AlaMetValCysThrArgArgAspGlnHisAlaLysSerValPheTyrIleAlaAlaGln  99

Qy    406  TGCCTGGGGCCCATCATTGGAGACAGGAAATCCTCATCGTGCACCTCCCAGTGTGGTG  465
           |||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db    100  CysLeuGlyAlaIleIleGlyAspArgLysSerSerValThrProProSerValVal  119

Qy    466  GGAGGCCTGGGAGTGCACCATGTTCATGGAAATCTTACCGCTGGTCTCTCTGGTT  525
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120  GlyGlyLeuGlyValThrThrValHisGlyValHisArgLeuLeuLeuLeuVal  139

Qy    526  GAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCCAAACGG  585
           |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    140  GluLeuIleIleThrPheGlnLeuValPheThrValPheAlaSerCysAspSerLysArg  159

Qy    586  ACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTCTGTTGCAATTGGACATTTA  645
           ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db    160  ThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeu  179

Qy    646  TTTGCAATCAATTATACTGTGCCAGCACATGAAATCCGCCGATCCTTTGACCTGCAGTT  705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180  PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaAspArgSerPheGlyAlaVal  199

Qy    706  ATCATGGGAAAATTGGGAGAAAACCATTGGATATATTGGGGTTGGGCCCATCATAGGAGCTGTC  765
           ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db    200  IleMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGlyAlaVal  219

Qy    766  CTCGCTGGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTGAATTCAAACGTCGTTTT  825
           ||||||||||||||||||||||||||:||||||||||||||||||||||| ||||||
Db    220  LeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValAsnSerAsnValValLeu  239

Qy    826  AAAGAAGCCTTCAGCAAAGCTGCCCAGCAACAAAAGGAAGCTACATGGAAGGTGGAGGAC  885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240  LysGluAlaPheSerLysAlaAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAsp  259

Qy    886  AACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATT  945
           |||||||||||||||:::|||||||||||||||||:||||||||||||||||||||||
Db    260  AsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHisValIle  279

Qy    946  GACGTTGACCGGGAGGAGAAGAAGAGGGGAAAGACCAATCTGGAGAGGTATTGTCTTCA  1005
           ||||:::||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    280  AspIleAspArgGluGluLysLysArgGlyLysThrAsnLeuGluArgSerGlyValLeuSerSer  299

Qy   1006  GTA   1008
           |||
Db    300  Val   300
```

RESULT 13
US-10-732-923-15746
; Sequence 15746, Application US/10732923
; Publication No. US2005010879lAl
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15746

(left column)

```
Db         |||||||||||||       |||||||
      121  GlyGlyLeuGlyValThrThrValHisGlyValHisGlyLeuLeuVal  140

Qy    526  GAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCCAAACGG  585
           |||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db    141  GluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArg  160

Qy    586  ACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTCTGTTGCAATTGGACATTTA  645
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161  ThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeu  180

Qy    646  TTTGCAATCAATTATACTGTGCCAGCACATGAATCCGCCGATCCTTTGACCTGCAGTT  705
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaAspArgSerPheGlyAlaVal  200

Qy    706  ATCATGGGAAAATTGGGATATATTGGGGTTGGGCCCATCATAGGAGCTGTC  765
           ||||||||||||||||:::|||||||||||||||||||||||||||||||
Db    201  IleMetGlyLysAsnTrpAlaAsnHisHisTrpIleTyrIleGlyValGlyProIleMetGlyAlaVal  220

Qy    766  CTCGCTGGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTGAATTCAAACGTCGTTTT  825
           ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db    221  LeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValAsnSerAsnValValLeu  240

Qy    826  AAAGAAGCCTCAGCAAAGCTGCCCAGCAAACAAAGAAGGAAGCTACATGGAAGGTCGGAGGAC  885
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  LysGluAlaPheSerLysAlaAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAspArgLeu  260

Qy    886  AACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATT  945
           |||||||||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db    261  AsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHisValIle  280

Qy    946  GACGTTGACCGGGAGGAGAAGAAGAGGGGAAAGACCAATCTGGAGAGGTATTGTCTTCA  1005
           ||||:::||||||||||||||||||||||||||:||||||||||||||||||||||
Db    281  AspIleAspArgGlyGluGluLysLysLysArgGlyLysLysAspSerSerGlyValGluValLeuSerSer  300

Qy   1006  GTA   1008
           |||
Db    301  Val   301
```

RESULT 12
US-10-732-923-15580
; Sequence 15580, Application US/10732923
; Publication No. US2005010879lAl
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15580
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-732-923-15580

Alignment Scores:
Pred. No.:   Length: 300
Score: 1.81e-132   Matches: 281
Percent Similarity: 1458.50   Conservative: 8
Best Local Similarity: 96.01%   Mismatches: 11
Query Match: 93.36%   Indels: 1
DB: 66.87%   Gaps: 1
                         17

US-10-723-180-1 (1-1152) x US-10-732-923-15580 (1-300)

```
Qy    106  ATGGTGGGCTTTCAAAGGGGTCTCGACTCAAGCTTTCTGAAAGCAGTCACAGCGGAATTT  165
           ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db      1  MetValAlaPheLysGlyValTrpThrGlnAlaPheThrGlnAlaValSerAlaValGluPhe  20

Qy    166  CTGGCCATGTCTTATTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACA  225
```

```
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Aedes aegypti
US-10-732-923-15746

Alignment Scores:
Pred. No.:              1.57e-54
Score:                  656.50
Percent Similarity:     69.72%
Best Local Similarity:  54.13%
Query Match:            30.10%
DB:                     17

Length:        249
Matches:       118
Conservative:  34
Mismatches:    63
Indels:        3
Gaps:          1

US-10-723-180-1 (1-1152) x US-10-732-923-15746 (1-249)

Qy  142  TGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCTCCTCAGCCTGGGA  201
              ||||:: ::  |||||||||||        :::  |||  |||
Db   23  TrpArgMetLeuValAlaGluPheLeuGlyThrPhePheLeuValSerIleGlyIleGly  42

Qy  202  TCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCC  261
              ||||||:::  |||||  |||:::          ||| :::
Db   43  SerThrMetGlyTrpGlyTyrArgLysAspTyrAlaPro--------ThrMetThrGlnIleAla  59

Qy  262  CTTTGCTTTGGACTCAGCATTGCAACCATGGCCATATCAGCGGTGGC  321
              |||||||||||  |||||||| ::  |||||||||||||
Db   60  PheThrPheGlyLeuValValAlaThrLeuAlaGlnAlaPheGlyHisValSerGlyCys  79

Qy  322  CACATCAACCCTGCAGTGACTGTGCCCATGGTGCCACCAGGAGATCAGCATCGCCAAG  381
              |||||||| |||||||||||   :::::  |||||||||||
Db   80  HisIleAsnProAlaValThrIleGlyLeuMetIleThrAlaAspIleSerIleLeuLys  99

Qy  382  TCTGTCTTCTACATCGCAGCCCAGTGCTGGGGGCCATCATTGGAGCAGGAATCCTCTAT  441
              |||||||||||  :::||||||:::  |||||||||||||    :::::
Db  100  GlyAlaPheTyrIleValSerGlnCysValGlyValAlaIleAlaAlaLeuValIleLys  119

Qy  442  CTGGTCACACCTCCACAGTGGTGGGAGGCCTGGAGTCACCATGGTTCATGGAAATCTT  501
              |||||||||||  :::||||||||||||  |||||  |||||||:::
Db  120  AlaAlaThrProSerAspValIleGlyLeuGlyValThrGlyLeuGlyIleAspProArgLeu  139

Qy  502  ACCGCTGGTCATGTCTCCTGGTTGAGTTGATAATCACATTTCAATTGTGTTTACTATC  561
              |||:::  ||||  :::||||||||||||||||||  ||||||
Db  140  ThrAlaGlyValMetGlyValMetGluAlaLeuGlyValLeuIleLeuThrPheLeuValVal  159

Qy  562  TTTGCCAGCTGTGATTCCAAACGGACTGATGTCCACTGGCTCAATAGCTTTAGCAATTGGA  621
              ||||||||  ||||||||:::::  |||||||||||||||||||
Db  160  HisGlyValCysValMetGlyAspAsnArgArgSerAspIleLysIleLeuGlySerAlaProLeuAlaIleGly  179

Qy  622  TTTTCTGTTGCAATTGGACATTATTGCAATCAATTATACTGGTGCCAGCATGAATCCC  681
              ::: |||||||||  ||||||    |||||||||||||  |||||
Db  180  LeuSerIleThrAlaGlyHisAlaGlyHisLeuSerAlaIleLeuLysTyrTyrThrGlyAlaSerMetAsnPro  199

Qy  682  GCCCGATCCTTTGACCTGCAGTTATCATGGAAAAACCATTGGATATATTGG  741
              |||||||  :::|||||||||||||||:::  |||||||||||
Db  200  AlaArgSerPheGlyProAlaValAlaValMetGlyValMetGlyValTrpThrAspGlnTrpTrpValTyrTrp  219

Qy  742  GTTGGGCCCATCATAGGAGAGCTGTCCTCGCTGGTGGCCTTTATGAGTATGTCTTC  795
              :::||||||  :::|||||||||||||||||||||
Db  220  ValGlyProIleValGlyIleValLeuAlaLeuAlaGlyLeuIleTyrArgLeuPhePhe  237
```

```
;   LENGTH: 251
;   TYPE: PRT
;   ORGANISM: Anopheles gambiae str. PEST
US-10-732-923-15615

Alignment Scores:
Pred. No.:              1.82e-51
Score:                  625.00
Percent Similarity:     67.43%
Best Local Similarity:  52.29%
Query Match:            28.66%
DB:                     17

Length:        251
Matches:       114
Conservative:  33
Mismatches:    67
Indels:        4
Gaps:          1

US-10-723-180-1 (1-1152) x US-10-732-923-15615 (1-251)

Qy  142  TGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCTCCTCAGCCTGGGA  201
              ||||:: :::  |||||||||||        :::  |||  ::: ||
Db   24  TrpArgMetLeuValAlaGluPheLeuGlyThrPhePheLeuValAlaIleGlyIleGly  43

Qy  202  TCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCC  261
              ||||||  |||||  |||:::            ||
Db   44  SerThrThrGlyTrpTyrThrAspTyr:SerProThrLeu--------ThrGlnIleAla  59

Qy  262  CTTTGCTTTGGACTCAGCATTGCAACCATGGCCATATCAGCGGTGGC  321
              |||||||||||  :::||||||  ::  |||||||||||||
Db   60  PheThrPheGlyLeuValValAlaThrLeuAlaGlnAlaPheGlyHisValSerGlyCys  79

Qy  322  CACATCAACCCTGCAGTGACTGTGCCCATGGTGCCACCAGGAGATCAGCATCGCCAAG  381
              ||||||||||||||||||||   ::::::  ::::::
Db   80  HisIleAsnProAlaValThrLeuGlyLeuLeuValLeuThrAlaAspValSerIleLeuLys  99

Qy  382  TCTGTCTTCTACATCGCAGCCCAGTGCTGGGGGCCATCATTGGAGTCAGGAATCCTCTAT  441
              |||||||||||  :::||||||:::  |||||||||||||    ::::::
Db  100  GlyAlaPheTyrIleValSerGlnCysIleSerGlyValAlaIleIleGlyAlaLeuValIleLys  119

Qy  442  CTGGTCACACCTCCACAGTGGTGGGAGGCCTGGAGTCACCATGGTTCATGGAAATCTT  501
              |||||||||||  :::||||||||||||  ::::::  |||||||||
Db  120  AlaAlaThrProSerGluValValGlyValAlaGlyLeuValLeuThrGlyLeuGlyValLeuIleLeuAlaProGlyLeu  139

Qy  502  ACCGCTGGTCATGTCTCCTGGTTGAGTTGATAATCACATTTCAATTGTGTTTACTATC  561
              :::  ||||||  :::||||||||||||||||||||||||
Db  140  SerThrGlyValGlnGlyValValLeuLeuGluLeuAlaLeuLeuThrPheMetLeuValPheVal  159

Qy  562  TTTGCCAGCTGTGATTCCAAACGGACTGATGTCCACTGGCTCAATAGCTTTAGCAATTGGA  621
              :::||||||  ||||||::: ||||||||||||||||||
Db  160  HisGlyValValCysCysAspAsnArgArgSerGlySerAlaProValLeuAlaIleGly  179

Qy  622  TTTTCTGTTGCAATTGGACATTATTGCAATCAATTATACTGGTGCCAGCATGAATCCC  681
              :::  |||||||||  ||||||    :::|||||||||||||  |||||
Db  180  LeuSerIleThrAlaGlyHisAlaGlyHisAlaAlaIleLeuLysTyrThrGlyAlaSerMetAsnPro  199

Qy  682  GCCCGATCCTTTGACCTGCAGTTATCATGGAAAAACCATTGGATATATTGG  741
              |||||||  :::|||||||||||||||:::  ||||||||||
Db  200  AlaAlaArgSerPheGlyProAlaValAlaValMetGlyValMetGlyValTrpThrAspGlnTrpTrpValTyrTrp  219

Qy  742  GTTGGGCCCATCATAGGAGAGCTGTCCTCGCTGGTGGCCTTTATGAGTATGTCTTC  795
              :::||||||  :::|||||||||||||||||||||
Db  220  ValGlyProIleValGlyIleValLeuAlaLeuAlaGlyLeuIleTyrArgLeuPhePhe  237
```

; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN

```
; APPLICANT:  Peter VEIBY
; APPLICANT:  Gordon B. MILLS
; APPLICANT:  Robert C. BAST, Jr.
; APPLICANT:  Karen LU
; APPLICANT:  Rosemarie SCHMANDT
; APPLICANT:  Xumei ZHAO
; APPLICANT:  Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;    LENGTH: 265
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-097-340-12
```

```
Alignment Scores:
Pred. No.:                      1.11e-50          Length:            265
Score:                          617.00            Matches:           132
Percent Similarity:             64.64%            Conservative:      38
Best Local Similarity:          50.19%            Mismatches:        81
Query Match:                    28.29%            Indels:            12
DB:                             14                Gaps:              4
```

US-10-723-180-1 (1-1152) x US-10-097-340-12 (1-265)

```
Qy   118 AAAGGGGTCTGGACTTCAAGCTTTCTGGAAAGCAGTCACAGCGGGAATTTCTGGCCATGCTT 177
            |||        |:    |||||||||| |||||||||||||||  ||||||||||
Db     3 LysGluValCysSerValAlaPheLeuLysGluAlaValPheAlaValPheLeuAlaThrLeu 22

Qy   178 ATTTTTGTTCTTCCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTA 237
            |||||||||                    |||  :::|||||
Db    23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp--------ProSerAlaLeu 39

Qy   238 CCGGTCGAACATGGTTCTCATCTCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAG 297
            ||:     |||||||||||:||||  |||||||:::|||   |||:|
Db    40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

Qy   298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTAGACCCTGTGGCCATGGTGTGC 357
            |||            ::|||||||||||||:||||||||||||| :::||||||||
Db    59 AlaLeuGlyProValSerGlyValGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuVal 78

Qy   358 ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCC 417
            :::||||||:::    :::::|  ||||||:::||||||    ::||||||
Db    79 GlyAsnGlnIleSerLeuArgLeuArgAlaPhePheTyrValAlaAlaGlnLeuValGlyAla 98

Qy   418 ATCATTGGAGCAGGAGAATCCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGGCCTGGGA 477
            ||||||||||||||||||  ::    |||  |||||||||||||||||||||
Db    99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaLeuProLeuGlyValAlaArgGlyAsnLeuAla 118

Qy   478 GTCACCATGGTTCATGGAAATCTTACCGCGTGGTCATGGTCTCCTGGTTGATAATC 537
            |||       |||      ||| |||     :::||||||||||||
Db   119 ValAsnAlaLeuAsnAsnThrThrGlnAlaMetValGlnAlaMetValGluLeuThrLeuLeu 138

Qy   538 ACATTCAATTGGTGTTTACTATCTTTGCCAGCTGTGAATTCCAAACGTCGATCGTCACT 597
```

```
Db   139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgArgThrSerProVal 158
            |||||||||                      |||||||||::|||||

Qy   598 GGCTCAATAGCTTAGCAGTTGGATTTCTGTTGCAATTGACAACATTTATTTGCAATCAAT 657
            ||||||  |||||:||||||||||||||||||||||||||  :::|||||||

Db   159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy   658 TATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAAT 717
            ::|||||||     |||||||||||||||||||||||||||||||     :::|||

Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198

Qy   718 TGGGAA---AACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCTGGT 774
            :::        ||||||||||||||||:::|||||||::||||||||

Db   199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla 218

Qy   775 GGCCTTTATGAGTAGTGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
            ||||||||||||   |||     |||::|||

Db   219 IleLeuTyrPheTyrLeuLeuPheProAsn------------SerLeuSer 231

Qy   835 TTCAGCAAAGCTGCCCAGCAAACAAAAGGAGCTACATGGAGGTGGAGGACAACAGGAGT 894
            ||||::|||||              ||||:::||||::||||||

Db   232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspTrpGluGlu 251

Qy   895 CAGGTAGAG 903
            ||| |||
Db   252 GlnArgGlu 254
```

Search completed: September 8, 2005, 12:23:27
Job time : 200 secs

This Page Blank (uspto)

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:    September 8, 2005, 11:55:58 ; Search time 50.5 Seconds
                                    (without alignments)
                             4389.771 Million cell updates/sec

Title:     US-10-723-180-1
Perfect score: 2181
Sequence:  1 ggggcaggcaatgagagctg.......attaccagtctaaacaata 1152

Scoring table: BLOSUM62
        Xgapop  10.0 , Xgapext   0.5
        Ygapop  10.0 , Ygapext   0.5
        Fgapop   6.0 , Fgapext   7.0
        Delop    6.0 , Delext    7.0

Searched:   283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:     566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  PIR_79:*
        1: pir1:*
        2: pir2:*
        3: pir3:*
        4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES
```

```
                      %
Result          Query
NO.    Score    Match Length DB  ID        Description
------ -----    ----- ------ --  --        -----------
  1  1689.5     77.5    341   2  I39178    aquaporin 4, long
  2  1595       73.1    323   2  I59283    water channel prot
  3   614       28.2    265   2  A55630    aquaporin-5, saliv
  4   612       28.1    263   2  A55279    major intrinsic pr
  5   606       27.8    263   1  MMBOLM    lens fiber membran
  6   602       27.6    262   2  JN0557    lens fiber membran
  7   600       27.5    261   2  S53423    major intrinsic pr
  8   597       27.4    271   2  A53442    aquaporin 2 - huma
  9   594       27.2    271   2  JT0750    water channel prot
 10   589.5     27.0    272   2  I51164    chip aquaporin - e
 11   589       27.0    271   2  I64818    water-channel aqua
 12   573       26.3    269   2  B44499    major intrinsic pr
 13   572       26.2    269   2  JC1320    water channel prot
 14   569       26.1    271   2  JC2348    water channel prot
```

```
15   563      25.8              269   2   I52366   uterine water chan
16   561      25.7              269   2   A41616   erythrocyte integr
17   443.5    20.3              278   2   A84545   hypothetical prote
18   443.5    20.3              309   2   T14889   membrane intrinsic
19   439.5    20.2              286   2   T14601   plasma membrane ma
20   437      20.0              285   2   T12435   probable plasma me
21   436      20.0              286   2   T02451   plasma membrane in
22   434      19.9              286   2   T47935   plasma membrane in
23   434      19.9              290   2   T02879   probable plasma me
24   433.5    19.9              283   2   T12434   probable plasma me
25   433      19.9              286   2   A86147   hypothetical prote
26   430.5    19.7              281   2   T14600   plasma membrane ma
27   430      19.7              295   2   T01528   probable plasma me
28   429.5    19.7              281   2   T09124   probable aquaporin
29   429.5    19.7              287   2   T12440   mipC protein - com
30   427.5    19.6              280   2   T05780   plasma membrane in
31   425.5    19.5              248   2   T12632   water channel prot
32   425      19.5              284   2   T12557   mipE protein - com
33   425      19.5              286   2   S42542   ripening-associate
34   424      19.4              288   2   T09794   major intrinsic pr
35   418.5    19.2              250   2   T14000   aquaporin TIP7 - C
36   418      19.2              287   2   T09791   drought-induced ma
37   418      19.2              700   2   S09699   bib protein - frui
38   415.5    19.1              288   2   T14863   porin Mip1 - Norwa
39   415      19.0              415   2   S41194   transmembrane prot
40   414.5    19.0              248   2   T14001   aquaporin TIP18 -
41   413      18.9              286   2   T04367   plasma membrane in
42   412.5    18.9              248   2   T10804   tonoplast intrinsi
43   412.5    18.9              288   2   T14599   plasma membrane ma
44   412      18.9              287   2   T05378   probable plasma me
45   410.5    18.8              285   2   E84789   hypothetical prote
```

```
                            ALIGNMENTS

RESULT 1
I39178
aquaporin 4, long splice form - human
N;Alternate names: mercurial-insensitive water channel protein; MIWC protein
N;Contains: aquaporin 4, long splice form; aquaporin 4, medium splice form; aquaporin 4,
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence revision 24-May-1996 #text_change 09-Jul-2004
C;Accession: I39178; I39177; JC6162; S66273
R;Yang, B.; Ma, T.; Verkman, A.S.
J. Biol. Chem. 270, 22907-22913, 1995
A;Title: cDNA cloning, gene organization, and chromosomal localization of a human mercu
A;Reference number: I39177; MUID:96032721; PMID:7559426
A;Accession: I39178
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-341 <YAN1>
A;Cross-references: UNIPROT:P55087; EMBL:U34846; NID:g1072054; PIDN:AAC52112.1; PID:g107
A;Note: alternatively spliced long form with upstream start codon
A;Accession: I39177
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 41-341 <YAN2>
A;Cross-references: EMBL:U34845; NID:g1072045; NID:g1072052; PIDN:AAC50284.1; PID:g1072053
R;Lu, M.; Lee, M.D.; Smith, B.L.; Jung, J.S.; Agre, P.; Verdijk, M.A.J.; Merkx, G.; Rijt
Proc. Natl. Acad. Sci. U.S.A. 93, 10908-10912, 1996
A;Title: The human AQP4 gene: Definition of the locus encoding two water channel polype
A;Reference number: JC6162; MUID:97008105; PMID:8855281
A;Accession: JC6162
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 19-263, 'G', 265-304, 'VE', 307-313, 'P', 315-341 <LUA>
A;Cross-references: GB:U63622; NID:g1680707; PIDN:AAB26957.1; PID:g1680708
R;Misaka, T.; Abe, K.; Iwabuchi, K.; Kusakabe, Y.; Ichinose, M.; Miki, K.; Emori, Y.; A
FEBS Lett. 381, 208-212, 1996
A;Title: A water channel closely related to rat brain aquaporin 4 is expressed in acid-
A;Reference number: S66273; MUID:9617?6324; PMID:8601457
A;Accession: S66273
A;Status: nucleic acid sequence not shown
```

A;Molecule type: mRNA
A;Residues: 19-263,'G',265-304,'VE',307-313,'P',315-341 <MIS>
A;Cross-references: GB:D63412; NID:g1236245; PIDN:BAA09715.1; PID:g1236246
C;Comment: This protein plays a role in the pathophysiology of normal pressure hydroceph
C;Genetics:
A;Gene: GDB:AQP4; MIWC
A;Cross-references: GDB:371722; OMIM:600308
A;Map position: 18q11.2-18q12.1
A;Introns: 29/2; 167/3; 222/3; 249/3
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: alternative initiators; alternative splicing; glycoprotein; transmembrane pr
F;1-341/Product: aquaporin 4, long splice form #status predicted <MAT1>
F;19-341/Product: aquaporin 4, medium splice form #status predicted <MAT2>
F;41-341/Product: aquaporin 4, short splice form #status predicted <MAT3>
F;56-71/Domain: transmembrane #status predicted <TM1>
F;88-111/Domain: transmembrane #status predicted <TM2>
F;115-117/Region: NPA motif
F;133-154/Domain: transmembrane #status predicted <TM3>
F;176-195/Domain: transmembrane #status predicted <TM4>
F;204-224/Domain: transmembrane #status predicted <TM5>
F;231-233/Region: NPA motif
F;250-270/Domain: transmembrane #status predicted <TM6>
F;171,224,301/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:   6.29e-133   Length: 341
Score: 1689.50   Matches: 326
Percent Similarity: 97.04%   Conservative: 2
Best Local Similarity: 96.45%   Mismatches: 5
Query Match: 77.46%   Indels: 5
DB: Gaps: 1

US-10-723-180-1 (1-1152) x I39178 (1-341)

Qy   10 AATGAGAGCTGC---------ACTCTGGCTCGGGAAGGCATGAGTGACAGACCC   54
          |||| |||:::||||||||||||||||||||
Db   4 AsnHisAlaCysPheValGluThrProAsnLeuAlaGlyGluGlyMetSerAspArgPro   23

Qy   55 ACAGCAAGGCGGTGGGTAAGTGGGACCTTTGTGTACCAGAGAGAACATCATGGTGGCT   114
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   24 ThrAlaArgArgTrpGlyLysTrpGlyProLeuCysTyrThrArgGluAsnIleMetValAla   43

Qy   115 TTCAAAGGGGTCTCGACTCCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTCTCGGCCATG   174
          |||||||||||||||||||||||||||||||||||||||||||||
Db   44 PheLysGlyValTrpThrProThrArgProGlnAlaPheTrpLysAlaValThrAlaGluPheLeuAlaMet   63

Qy   175 CTTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCT   234
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 LeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlyThrGluLysPro   83

Qy   235 TTACCGGTCGACATGGTTCTCATCCCCCTTGCTTGGACTCAGCATTGCAACCATGGTG   294
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   84 LeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetVal   103

Qy   295 CAGTGCTTTGGCCATATCAGCGGTGCCACATCAACCCTGCAGTGACTGTGGCCATGGTG   354
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   104 GlnCysPheGlyHisIleSerGlyAlaHisIleAsnProAlaValThrValAlaMetVal   123

Qy   355 TGCACCAGGAAGATCAGCATCGCCAGTCTGTCTTCACATCGAGCCCAGTGCTCTGGGG   414
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 CysThrArgLysIleSerIleAlaSerLeuSerSerValPheTyrIleAlaAlaGlnCysLeuGly   143

Qy   415 GCCATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTG   474
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   144 AlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSerValValGlyGlyLeu   163

Qy   475 GGAGTGCCACCATGGTTCATGAAAATCTTACCGCTGCTCTCCTGGTTGAGTTGATA   534
          ||||||||||||||||||||||||||||||||||||||||||||
Db   164 GlyValThrMetValHisGlyAsnLeuThrAlaAlaGlyHisGlyLeuLeuValGluLeuIle   183

Qy   535 ATCACATTTCAATTGGTGTTACTATCTTGCCAGCTGTGATTCCAAACGGACTGATGTC   594
          ||||||||||||||||||||||||||||||||||||||||||||
Db   184 IleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgThrAspVal   203

Qy   595 ACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTGCAATC   654
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   204 ThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIle   223

Qy   655 AATTATACTGGTGCCAGCATGAATCCGCCCGATCCTTGGACCTGCAGTTATCATGGGA   714
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   224 AsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetGly   243

Qy   715 AATTGGGAAAACCATTGGATAATATGGGTTGGGCCCATCATAGGAGCGTGTCCTCGCTGGT   774
          |||||||||||||||||||||||||||||||||||||||||||||
Db   244 AsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaValAlaLeuAlaGly   263

Qy   775 GGCCTTATGAGTAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGCC   834
          |||||||||||||||||||||||||||||||||||||||||||||
Db   264 AlaLeuTyrGluTyrValPheCysProAspValGluPheLysSerArgArgPheLysGluAla   283

Qy   835 TTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAAGTGGAGACAACAGGAGT   894
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   284 PheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluThrThrGlySer   303

Qy   895 CAGGTAGACAGGATGACCTGATTCTAAAACCTGAGTGGTGCATGTGATTGACCTTGAC   954
          ||| :::||||||||||||||||||||||||||||||||||||
Db   304 GlnValAspArgMetThrLeuIleLeuLysProGluTrpTrpHisValHisAspValAspAsp   323

Qy   955 CGGGGAGGAGGAGAAGAAGGGGAAAGACCAATCTGAGAGGTATTGTCTTCAGTA   1008
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   324 ArgGlyGlyGlyGluLysLysGlyLysAspGlnSerGlyGluValLeuSerSerVal   341

RESULT 2
I59283
water channel protein, mercurial-insensitive - rat
N;Alternate names: aquaporin 4; MIWC
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I59283; A53194
R;Jung, J.S.; Bhat, R.V.; Preston, G.M.; Guggino, W.B.; Baraban, J.M.; Agre, P.
Proc. Natl. Acad. Sci. U.S.A. 91, 13052-13056, 1994
A;Title: Molecular characterization of an aquaporin cDNA from brain: candidate osmorecep
A;Reference number: I59283; MUID:95108097; PMID:7528931
A;Accession: I59283
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-323 <RES>
A;Cross-references: UNIPROT:P47863; EMBL:U14007; NID:g595787; PIDN:AAC52152.1; PID:g5957
A;Experimental source: brain
R;Hasegawa, H.; Ma, T.; Skach, W.; Matthay, M.A.; Verkman, A.S.
J. Biol. Chem. 269, 5497-5500, 1994
A;Title: Molecular cloning of a mercurial-insensitive water channel expressed in selecte
A;Reference number: A53194; MUID:94164885; PMID:7509789
A;Accession: A53194
A;Molecule type: mRNA
A;Residues: 23-200,'P',202-323 <HAS>
A;Cross-references: GB:L27588; NID:g459950; PIDN:AAA17730.1; PID:g459951
C;Genetics:
A;Gene: AQP4
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: glycoprotein; phosphoprotein; transmembrane protein
F;30-57/Domain: transmembrane #status predicted <TMM1>
F;70-89/Domain: transmembrane #status predicted <TMM2>
F;115-136/Domain: transmembrane #status predicted <TMM3>
F;159-177/Domain: transmembrane #status predicted <TMM4>
F;185-205/Domain: transmembrane #status predicted <TMM5>
F;231-252/Domain: transmembrane #status predicted <TMM6>
F;107/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;153/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;180/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Alignment Scores:
Pred. No.:   Length: 323
Score: 4.84e-125   Matches: 304
Percent Similarity: 96.59%   Conservative: 8
Best Local Similarity: 94.12%   Mismatches: 11

```
Query Match:      73.13%        Indels:       0
DB:               2            Gaps:          0

US-10-723-180-1 (1-1152) x I59283 (1-323)

Qy    40 ATGAGTGACAGACCCCAGCAAGGCGGTGGGTAAGTGTGACACCTTTGTGTACCAGAGAG  99
          ||||||||||||||||||||||||||||||||||||||||||||  :::||||||
Db     1 MetSerAspGlyAlaAlaAlaAlaArgArgTrpGlyLysValTrpCysGlyProCysSerArgGlu  20

Qy   100 AACATCATGTGGCCTCTGACTCCTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCG 159
          :::||||||||||||||||||||||||||||||||||||||||||||||
Db    21 SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla  40

Qy   160 GAATTTCTGGCCATGCTATTTTGTCTCCTCAGCTGGGATCCACCATCAACTGGGGT 219
          |||||||||||||||||||||||||||  :::|||||||||||||||
Db    41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerValGlySerThrIleAsnTrpGly  60

Qy   220 GGAACAGAAAAGCCTTACCGGTCGACATGGTTTCTCATCTCCCTTTGCTTGGACTCAGC 279
          |||:::|||||||||||||||||||
Db    61 GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer  80

Qy   280 ATTGCAACCATGGTGCAGTGCTTTGGCCATATCAGCGGTGGCCACCATCAACCCTGCAGTG 339
          |||||||||||||||||||||||||||||||||||||
Db    81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisHisGlnProAlaVal 100

Qy   340 ACTGTGGCCATGGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA 399
          |||||||||||||||||||||||||||||||||||||||
Db   101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleThr 120

Qy   400 GCCCAGTGCCTGGGGCCCATCATTGGAGCAGGAATCCTCTATCTGGTCACCCTCCCAGT 459
          |||||||||||||||||||||||||||||||||||||
Db   121 AlaGlnCysLeuGlyProIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer 140

Qy   460 GTGGTGGGAGGGCCTGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC 519
          |||||||||||||||||||  ||||||||||||||||
Db   141 ValValGlyGlyLeuGlyValThrThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160

Qy   520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTTACTATCTTTGCCAGCTGTGATTCC 579
          ||||||||||||||||||||||||||||
Db   161 LeuValGluLeuIleIleThrPheAsnTrpCysLeuLeuSerLeuProAlaValIlePro 180

Qy   580 AAACGGACTGCTGTGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGA 639
          |||||||||||||||||:::||||||||||||||||
Db   181 LysArgThrAlaValLeuAsnSerPheSerAsnTrpIlePheCysCysAsnTrpIle 200

Qy   640 CATTTATTTGCAATCAATTATATCGGTGCCAGCATGAATCCCGCCCGATCCTTTGGACCT 699
          :::||||||||||||  :::|||||||||||
Db   201 HisLeuPheAlaIleAsnTyrIleGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220

Qy   700 GCAGTTATCATGGGAAAACCATTGGATATTGGGTTGGGCCCATCATAGGA 759
          |||||||||||||||  :::|||||||||||
Db   221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240

Qy   760 GCTGTCCTCGGTGGCCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGT 819
          ||||||||||||||||||||
Db   241 AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260

Qy   820 CGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGAGTACATGGAGGTG 879
          ||||||||||||:::|||||||||||
Db   261 ArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysSerTyrMetGluVal 280

Qy   880 GAGGACAACAGGAGTCAGGTAGACGATGACCTGATTCTAAAACCTGGAGTGGTGCAT 939
          ||||||||||:::||||||||||
Db   281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300

Qy   940 GTGATTGACGTTGACCCGGGAGAGGAGAAGACCAATCTGGAAGACTCTGGAGGTATTG 999
          |||||||||||:::|||||||||||
Db   301 ValIleAspValAspProGlyGluGluLysThrAsnLeuGluAspSerGlyGlyIleLeu 320

Qy  1000 TCTTCAGTA 1008
          |||||||||
Db   321 SerSerVal 323
```

RESULT 3
A55630
aquaporin-5, salivary gland - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
R;Raina, S.; Preston, G.M.; Guggino, W.B.; Agre, P.
J. Biol. Chem. 270, 1908-1912, 1995
A;Title: Molecular cloning and characterization of an Aquaporin cDNA from salivary, lacr
A;Reference number: A55630; MUID:95130575; PMID:7530250
A;Accession: A55630
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-265 <RAI>
A;Cross-references: UNIPROT:P47864; GB:U16245; NID:g664759; PIDN:AAA66221.1; PID:g664760
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: transmembrane protein

Alignment Scores:
Pred. No.:   Length: 265
Score: 614.00   Matches: 133
Percent Similarity: 56.83%   Conservative: 46
Best Local Similarity: 42.22%   Mismatches: 82
Query Match: 28.15%   Indels: 54
DB: 2   Gaps: 5

```
US-10-723-180-1 (1-1152) x A55630 (1-265)

Qy   118 AAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACACGGGAATTTCTGGCCATGCTT 177
          |||                :::                ||||||||
Db     3 LysGluValCysSerLeuAlaPheLysAlaValPheLysAlaValPheLysAlaValPheAlaProPheLysGluProLysAlaValPheAlaGluPheLeuAlaMetLeu  22

Qy   178 ATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTA 237
          ||||||||||||||               :::                |||
Db    23 IlePheValPheGlyLeuGlySerAlaLeuLysTrp------------ProSerAlaLeu  39

Qy   238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTGGACTCAGCTGTCAACCATGGTGCAG 297
          |||                :::                ||                :::
Db    40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuGlyLeuIleGlyThrAlaGln  58

Qy   298 TGCTTTGGCCATATCAGCGGTGGCCACCATCAACCCTGCAGTGACTGTGGCCATGGTGTGC 357
          ||                :::                ||
Db    59 AlaLeuGlyProValSerGlyGlyHisHisGlnProAlaIleThrValAlaLeuLeuLeu  78

Qy   358 ACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCACAGTGCCTGGGGGCC 417
          :::||||||               ::||||||||                  :::||||||
Db    79 GlyAsnGlnIleSerLeuLeuArgAlaValAlaValPheTyrValAlaAlaGlnLeuValGlyValGlyAla  98

Qy   418 ATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCAGTGTGGTGGGAGGGCCTGGGA 477
          ||                :::                      ||
Db    99 IleAlaGlyAlaGlyIleIleLeuTyrLeuValTrpLeuAlaProLeuAsnAlaArgGlyAsnAlaLeuAla 118

Qy   478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTTCCTCGGTTGAGTTGATAATC 537
          ||||||               :::                  ||
Db   119 ValAsnAlaLeuAsnAsnAsnThrArgProGlyLysGlyAlaMetValGluLeuIleLeu 138

Qy   538 ACATTCAATTGGTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGATGTCACT 597
          :::||||||               ::||||||
Db   139 ThrPheGlnLeuAlaLeuCysLeuLeuPheSerSerThrAspSerArgArgThrSerProVal 158

Qy   598 GGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTGCAATCAAT 657
          ||||||               :::                |||||||
Db   159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178

Qy   658 TATATCGGTGCCAGCATGAATCCCGCCCGATCCTTTGGACCTGCAGTTATCATGGGAAAT 717
          :::                |||||||
Db   179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198

Qy   718 TGGGAA---AACCATTGGATATATTGGGTGGGGCCCATCATAGGAGCTGTCCTCGGT 774
          :::                |||||||
Db   199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAlaAla 218

Qy   775 GGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAAGCC 834
```

```
Db   219 IleLeuTyrPheTyrLeuLeuPheProSer-----------------SerLeuSer 231
             ||||||                                      :::
Qy   835 TTCAGCAAAGCTGCCCAGCAAACAAAGGAGGAGCTACATGGAGGTGGAGACAACAGGAGT 894
             |||:|||    ||||            |:||
Db   232 LeuHisAspArgArgValAlaValAlaValLysGlyThrTyrGluProGluGluAsp--------- 248
Qy   895 CAGGTAGAGACGGATGACCTGATTCTAAAAACCTGGAGTGGTGCATGTGATGATGACGTTGAC 954
Db   248 --------------------------------------------------------- 248
Qy   955 CGGGGAGAGAGAAGAAGGGGAAAGACCAATCTGGAGAGGTATTGTCTTCAGTATGACTA 1014
Db   249 --------------------------------------------------------Trp 249
Qy  1015 GAAGATCGCACTGAAAGCAGACAAGACTCCTTAGAACTGTCCTCA 1059
             |||    ||:|| :::::/||||||:|::
Db   250 GluAspHisArgGluGluArgLysLysThrIleGluLeuThrArgAla 264
```

RESULT 4
A55279
major intrinsic protein - human
N;Alternate names: MP26
C;Species: Homo sapiens (man)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C;Accession: A55279
R;Pisano, M.M.; Chepelinsky, A.B.
Genomics 11, 981-990, 1991
A;Title: Genomic cloning, complete nucleotide sequence, and structure of the human gene
A;Reference number: A55279; MUID:92147144; PMID:1845563
A;Accession: A55279
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <PIS>
A;Cross-references: UNIPROT:P30301; GDB:119394; OMIM:154050
A;Experimental source: lens
A;Note: sequence extracted from NCBI backbone (NCBIN:80218, NCBIP:80220)
C;Genetics:
A;Gene: GDB:MIP
A;Cross-references: GDB:119394; OMIM:154050
A;Map position: 12q13-12q13
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: eye lens; transmembrane protein

```
Alignment Scores:
Pred. No.:              5.46e-43      Length:        263
Score:                 612.00        Matches:       114
Percent Similarity:     65.27%       Conservative:   57
Best Local Similarity:  43.51%       Mismatches:     77
Query Match:            28.06%       Indels:         14
DB:                                  Gaps:            3

US-10-723-180-1 (1-1152) x A55279 (1-263)
```

```
Qy   124 GTCTGG-------ACTCAAGCTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATG 174
             ::/||         :::|||||||:::|||||||||    ||
Db     1 MetTrpGluLeuArgSerAlaSerPheTrpArgAlaIlePheAlaGluPheAlaThr 20
Qy   175 CTTATTTTTGTTCTCCTCAGCCTGGGATCCACCATCAACCTGGGTGGAACAGGAAAAGCCT 234
             ||| :::|||||||||||||||||||||:::||
Db    21 LeuPheTyrValPheGlyLeuGlySerSerArgTrpAla-------------Pro 36
Qy   235 TTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTGGACTCAGCATTGCAACCATGGTG 294
             |||   :::::   :::|||   ||||||||||||
Db    37 GlyProLeuHisValLeuGlnValAlaMetAlaPheGlyLeuAlaLeuAlaThrLeuVal 56
Qy   295 CAGTGCTTTGGCCATATCAGCGGTGGCCACATCAACCCCTGCAGTGACCTGTGGCCATGGTG 354
             ||||||    ||||||:::|||||:::|||||:::||||
Db    57 GlnSerValGlyHisIleSerGlyGlyHisIleValAlaAsnProAlaValThrPheAlaPheLeu 76
Qy   355 TGCACCAGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCCAGCCCAGTGCCTGGGG 414
             ||:::|| :::::: |||:::|||||||||||
Db    77 ValGlySerGlnMetSerLeuLeuArgArgAlaPheCysTyrMetAlaLeuAlaGlnLeuLeuGly 96
```

```
Db   415 GCCATCATTGGAGCAGGAATCCCTATCTGGTCACACCTTCCCAGTGCGTGGGAGAGGCCTG 474
             :::||||       |||||
Qy    97 AlaAlaAlaGlyAlaAlaValAlaLeuTyrSerValThrProProAlaValArgGlyAsnLeu 116
Db   475 GGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCACGGTGGTCCTCTGGTTGAGTTGATA 534
             :::       |||
Qy   117 AlaLeuAsnThrLeuHisProAlaValSerValGlyGlnAlaThrThrValGluIlePhe 136
Db   535 ATCACATTTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTC 594
             :::|||     ||||||:::
Qy   137 LeuThrLeuGlnPheValLeuCysIlePheAlaThrTyrAspGluArgArgAsnGlyGln 156
Db   595 ACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTTGCAATC 654
             :::           ||||||
Qy   157 LeuGlySerValAlaLeuAlaLeuAlaValGlyPheSerValAlaLeuGlyHisLeuPheGlyMet 176
Db   655 AATTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGA 714
             ||||||
Qy   177 TyrTyrThrGlyAlaGlyMetAsnProAlaArgSerPheAlaProAlaIleLeuThrGly 196
Db   715 AATTGGGAAAACCATTGGATATATATGGGTGGGCCCATCATAGAGCTGTCTCCTCGCTGGT 774
             ||||:::|||
Qy   197 AsnPheThrAsnHisTrpIleTyrTrpValTyrTrpValGlyProIleIleGlyGlyLeuGlySer 216
Db   775 GGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
             ||||||:::::::        |||
Qy   217 LeuLeuTyrAspPheLeuLeuPheProArgLeu-----------------LysSer 229
Db   835 TTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
             |||:|:        |||:||     ::|:|
Qy   230 IleSerGluArgLeuSerValLeuLysGlyAlaLeuAlaLysProAspValSerAsnGlyGlnPro 249
Db   895 CAGGTA 900
             ::|||
Qy   250 GluVal 251
```

RESULT 5
MMBOLM
lens fiber membrane major intrinsic protein - bovine
N;Alternate names: lens fiber membrane intrinsic 26K protein (MIP); MP26
N;Contains: MP22
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A23251; B23251; A22444
R;Gorin, M.B.; Yancey, S.B.; Cline, J.; Revel, J.P.; Horwitz, J.
Cell 39, 49-59, 1984
A;Title: The major intrinsic protein (MIP) of the bovine lens fiber membrane: characteri
A;Reference number: A23251; MUID:85024900; PMID:6207938
A;Accession: A23251
A;Molecule type: mRNA
A;Residues: 1-263 <GOR>
A;Cross-references: UNIPROT:P06624; GB:K02818; NID:g163296; PIDN:AAA30622.1; PID:g163297
A;Accession: B23251
A;Molecule type: protein
A;Residues: 1-39 <GO2>
R;Ngoc, L.D.; Paroutaud, P.; Dunia, I.; Benedetti, E.L.; Hoebeke, J.
FEBS Lett. 181, 74-78, 1985
A;Title: Sequence analysis of peptide fragments from the intrinsic membrane protein of c
A;Reference number: A22444; MUID:85127500; PMID:3882455
A;Accession: A22444
A;Molecule type: protein
A;Residues: 1-13,'L',15-28,'R',30,'F',32,'L',184-214 <NGO>
C;Comment: This protein is the major component of lens fiber gap junctions.
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: eye lens; gap junction; phosphoprotein; transmembrane protein
F;12-28/Domain: transmembrane #status predicted <TM1>
F;44-60/Domain: transmembrane #status predicted <TM2>
F;87-103/Domain: transmembrane #status predicted <TM3>
F;130-146/Domain: transmembrane #status predicted <TM4>
F;160-176/Domain: transmembrane #status predicted <TM5>

Alignment Scores:

Left column — RESULT 6 (MMBOLM)

```
Pred. No.:                        Length:        263
Score:                1.73e-42    Matches:       113
Percent Similarity:   64.07%      Conservative:  60
Best Local Similarity: 41.85%     Mismatches:    83
Query Match:          27.79%      Indels:        14
DB:                               Gaps:          3

US-10-723-180-1 (1-1152) x MMBOLM (1-263)

Qy  124  GTCTGG-------ACTCAAGCTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATG  174
              :::|||   ::|||||||::|||           |||||||
Db    1  MetTrpGluLeuArgSerAlaSerPheTrpArgAlaIleCysAlaPheAlaSer      20

Qy  175  CTTATTTTGTTCTCCTCAGCCTGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCT  234
         |||||||   ::|||  ||||||:::::::   |||   :::
Db   21  LeuPheTyrValPhePheGlyLeuGlyAlaSerLeuArgTrpAla---------Pro   36

Qy  235  TTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTG  294
              |||::  :::::   |||    ||| :::|||||::|||||||:::||
Db   37  GlyProLeuHisValLeuGlnValAlaLeuAlaPheGlyLeuLeuAlaLeuAlaThrLeuVal  56

Qy  295  CAGTGCTTTGGCCATATCAGCGCGGTGCCACATCAACCTGCAGTGACTGTGGCCATGGTG 354
         |||||  ||||||||   |||||:::   ||||| :::||||||  || :::
Db   57  GlnAlaValGlyHisHisIleSerGlyAlaHisValAsnProAlaValThrPheAlaPheLeu  76

Qy  355  TGCACCGGAAGATCAGCATCCGCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCTGGGG  414
             :::::|||:::   :::||||||:   ||||||  |||||:::||
Db   77  ValGlySerGlnMetSerLeuLeuArgAlaIleCysTyrMetValAlaGlnLeuLeuGly    96

Qy  415  GCCATCATTGGAGGAGGAATCCTCATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTG  474
         |||::   |||||||   ||||:::||||||||:::|||||
Db   97  AlaValAlaGlyAlaAlaValLeuTyrSerValThrProProAlaValArgGlyValAsnLeu 116

Qy  475  GGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTCCGTTGAGTTGATA    534
              :::   ||||||||   :::|||   :::|||||:::
Db  117  AlaLeuAsnThrLeuHisProGlyValSerValGlyGlnAlaThrIleValGluIlePhe   136

Qy  535  ATCCAATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTC    594
             ::|||   |||   ||||:::||||||:::
Db  137  LeuThrLeuGlnPheValLeuCysPheIlePheLeuGlyHisGlyLeu            156

Qy  595  ACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTGCAATTTATTTGCAAATC        654
         |||||||||   :::   ||||||||:::   :::
Db  157  LeuGlySerValAlaLeuAlaValAlaGlyPheMetAsnProAlaIleLeuValAla     176

Qy  655  AATTGGAAAAACCATTGGATTATTGGGTGGGCCCATCATAGAGAGCTGTCCTCGCTGGT 714
              |||||||   ::::|||   |||||::||||||||::::
Db  177  TyrTyrThrGlyAlaGlyMetAsnProAlaAlaLeuPheSerLeuTrpLeuThrLeuThrArg 196

Qy  715  AATTGGGAAAACCATTGGATGTATTATTGGGTGGGCCCATCATAGAGAGCTGTCCTCGT 774
         |||:::   :::|||||::||||||   |||||||::
Db  197  AsnPheThrAsnHisTrpValTyrTrpValGlyProValIleGlyAlaGlyLeuGlySer   216

Qy  775  GGGCCTTTATGAGTATGTCTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
              :::||||||||:::   |||||   ||||:::
Db  217  LeuLeuTyrAspPheLeuLeuPheProArgLeu-----------------LysSer     229

Qy  835  TTCAGCAAAGCTGCCCAAGTCGCCAGCAAACAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGT 894
         |||:::   :::
Db  230  ValSerGluArgLeuSerLeuLeuLysGlySerArgProSerGluSerAsnGlyGlnPro   249

Qy  895  CAGGTAGAGACGGATGACCTGATTCTAAAA                              924
         |||:::   ::: :::|||||||
Db  250  GluValThrGlyGluProValGluLeuLys                             259
```

```
RESULT 6
JN0557
lens fiber membrane major intrinsic protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 13-Aug-1999
C;Accession: JN0557
R;Austin, L.A.R.; Rice, S.J.; Baldo, G.J.; Lange, A.J.; Haspel, H.C.; Mathias, R.T.
Gene 124, 303-304, 1993
```

Right column — RESULT 7 (JN0557)

A;Title: The cDNA sequence encoding the major intrinsic protein of frog lens.
A;Reference number: JN0557; MUID:93185940; PMID:7916707
A;Accession: JN0557
A;Molecule type: mRNA
A;Residues: 1-262 <AUS>
A;Cross-references: EMBL:X56970; NID:g288514; PIDN:CAA40291.1; PID:g288515
A;Experimental source: lens
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: eye lens; gap junction; transmembrane protein

```
Alignment Scores:
Pred. No.:                        Length:        262
Score:                3.73e-42    Matches:       116
Percent Similarity:   65.60%      Conservative:  48
Best Local Similarity: 46.40%     Mismatches:    77
Query Match:          27.60%      Indels:        9
DB:                               Gaps:          2

US-10-723-180-1 (1-1152) x JN0557 (1-262)

Qy  136  GCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGGCTTATTTTGTTCTCCTCAGC 195
              :::|||||||::|||||||||||||||       ::: :::|||
Db    7  SerPheTrpArgAlaIlePheAlaGluPheGlyThrPheTyrValPhePheGly         26

Qy  196  CTGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTC   255
         ||||::|||   ||||:::|||       -----ProAlaAsnValLeuVal
Db   27  LeuGlyAlaSerLeuLysTrpAlaAlaGly--------ProAlaAsnValLeuVal      42

Qy  256  ATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCATATCAGC   315
         :::|||||||||:::   :::|||::|||||   |||
Db   43  IleAlaLeuAlaPheGlyLeuValLeuAlaThrMetValGlnSerIleGlyHisValSer   62

Qy  316  GCGGTGCCACATCAACCTGCAGTGACTGTGGCCATGGTGCACCGGAAGATCAGCATC     375
         |||   |||   |||:::||||| || :::
Db   63  GlyAlaHisIleAsnProAlaValPheAlaPheLeuIleGlySerGlnMetSerLeu     82

Qy  376  GCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCTGGGGGCCATCATTGGAGGAGGAATC  435
         :::   ::|||||||::: |||||:::||||||
Db   83  PheArgAlaIlePheTyrIleAlaAlaGlnLeuLeuGlyAlaValAlaValAlaAlaVal  102

Qy  436  CTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGA 495
         |||||||||   |||:::|||||||:::   :::||
Db  103  LeuTyrGlyValThrProAlaAlaIleThrArgValGluPheLeuHisPheValLeuHisPro 122

Qy  496  AATCTTACCGGTCGTCATGGTTCCTGGTGAGTTGAAATTGAATTCAATTTCAATTGGTGTT 555
         |||   :::|||   :::||
Db  123  GlyValSerLeuGlyGlnAlaThrValGluIlePheLeuThrLeuGlnPheValLeuVal   142

Qy  556  ACTATCTTTGCCAGCTGTGATTCAATTGGTGTCCATGATGCTGGCTCAATAGTCTTAGCA 615
         |||||||||:::   ||||||::::::
Db  143  CysIlePheLeuThrThrTyrAspGlyArgArgAsnGlyArgLeuGlySerValSerLeuAla 162

Qy  616  ATTGGATTTTCTGTGCAATTGGCAATTGGACATTATTGGCAATCAATTATACTGGTGCCAGCATG 675
         |||:::||||||   :::   :::||||||:::|||
Db  163  IleGlyPheSerLeuLeuThrLeuGlyHisLeuPheGlyLeuTyrTyrThrGlyAlaSerMet   182

Qy  676  AATCCCGGCCGATCCTTTGACCTGCAGTTATCATGGCCAGTTGGGAAAAACCATTGGATA 735
         |||:::|||   ::::::|||   :::||
Db  183  AsnProAlaArgSerPheAlaProAlaValLeuThrArgAsnPheThrAsnHisTrpVal   202

Qy  736  TATTGGGTTGGGCCCATCATAGAGAGCTGTCCTCGCTGGTGGCCCTTTATGAGTATGTCTTC 795
         |||:::   :::   :::|||   |||:::||
Db  203  TyrTrpValGlyProIleIleGlyAlaLeuGlyLeuGlyLeuValTyrAspPhePheIleLeu 222

Qy  796  TGTCCAGATGTT--------------------GAATTCAAACGTCGTTTTAAAGAAGCCTTCAGCAAA 843
              :::   ::: :::                :::
Db  223  PheProArgMetArgGlyLeuSerGlu--ArgLeuSerIleLeuLysGlyAlaArgProAl   242

Qy  844  GCTGCCCAGCAAACAAAGAAGCTACA                                    871
         |||   :::|||||||:::||
Db  242  aGluProGluGlyGlnGlnGluAlaThr                                  251
```

RESULT 7

S53423
major intrinsic protein (MIP26) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S53423; S10491
R;Dilsiz, N.; Crabbe, M.J.C.
Biochem. J. 305, 753-759, 1995
A;Title: Heterologous expression in Escherichia coli of native and mutant forms of the m
A;Reference number: S53423; MUID:95151007; PMID:7848273
A;Accession: S53423
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-261 <DIL>
A;Cross-references: UNIPROT:P09011
R;Kent, N.A.; Shiels, A.
Nucleic Acids Res. 18, 4256, 1990
A;Title: Nucleotide and derived amino-acid sequence of the major intrinsic protein of ra
A;Reference number: S10491; MUID:90332430; PMID:2377471
A;Accession: S10491
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-261 <KEN>
A;Cross-references: EMBL:X53052; NID:g57621; PIDN:CAA37219.1; PID:g1334235
C;Superfamily: lens fiber membrane major intrinsic protein

Alignment Scores:
Pred. No.:   261
Score: 600.00
Percent Similarity: 68.47%   Length: 261
Best Local Similarity: 48.65%   Matches: 108
Query Match: 27.51%   Conservative: 44
DB: Mismatches: 66
                                     Indels: 4
                                     Gaps: 2

US-10-723-180-1 (1-1152) x S53423 (1-261)

Qy   136 GCTTTCTGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTTATTTTTGTTCTCCTCAGC 195
Db   6 SerPheTrpArgAlaIlePheAlaGluPheAlaThrLeuPheTyrValPhePheGly 25

Qy   196 CTGGGATCACCACCAACTGGGGTGGAACAGAAAAGCCTTACCGGTCGACATGGTTCTC 255
Db   26 LeuGlySerSerLeuArgTrpAla---------ProGlyProLeuHisValLeuGln 41

Qy   256 ATCTCCCTTTGCTTTGGACTTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCATATCAGC 315
Db   42 ValAlaLeuAlaPheGlyLeuAlaLeuAlaThrLeuValGlnThrValGlyHisIleSer 61

Qy   316 GGTGGCCACCATCAACCCTGCAGTGACTGTGGCCATGGTGTCACCAGGAAGATCAGCATC 375
Db   62 GlyAlaHisValAsnProAlaValThrPheAlaPheLeuValGlySerGlnMetSerLeu 81

Qy   376 GCCAAGTCTGTCTTCTACATCGCCAGCCCAGTGCCTGGGGGCCATCATTGGAGCAGGAATC 435
Db   82 LeuArgAlaiPheCysTyrIleAlaSerAlaGlnLeuLeuGlyAlaValAlaGlyAlaVal 101

Qy   436 CTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTTGGGAGTCCACCATGGTTCATGGA 495
Db   102 LeuTyrSerValThrProProAlaValArgGlyAsnLeuAlaLeuAsnThrLeuHisAla 121

Qy   496 AATCTTACCCGCTGGTACATGTCTCCTGGTTGAGTTGAATTCAATTGGTGTTT 555
Db   122 GlyValSerValGlyGlnAlaThrThrValGluIlePheLeuThrLeuGlnPheValLeu 141

Qy   556 ACTATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGCTCAATAGCTTTAGCA 615
Db   142 CysIlePheAlaThrTyrAspGluArgArgMetGlyArgMetGlyLeuGlySerValAlaLeuAla 161

Qy   616 ATTGGATTTCTGTTGCAATTGACAATTTATTGCAATCAATTATACTGGTGCCAGCATG 675
Db   162 ValGlyPheSerLeuThrLeuGlyHisLeuPheGlyMetTyrTyrThrGlyAlaGlyMet 181

Qy   676 AATCCCGCCCGATCCTTTGACCTGCCAGTTATCATGGGAAATTGGGAAAACCATTGGATA 735
Db   182 IleAlaMetAlaPheGlyLeuValGlyIleGlyThrLeuuValGlnAlaLeuuGlyHisIleSer 63

Db   182 AsnProAlaAargSerPheAlaProAlaIleLeuThrArgAsnPheSerAsnHisTrpVal 201
Qy   736 TATTGGGTTGGGCCATCATAGGAGCTGTTCCTGCGTGGTGGCCTTTATGAGTATGTCTTC 795
Db   202 TyrTrpValGlyProIleIleGlyIleGlyLeuGlySerLeuuLeuTyrAspPheLeuuLeu 221

Qy   796 TGTCCA 801
Db   222 PhePro 223

RESULT 8
A53442
aquaporin 2 - human
N;Alternate names: water-channel protein
C;Species: Homo sapiens (man)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: A53442; I55410; I51877
R;Deen, P.M.T.; Verdijk, M.A.J.; Knoers, N.V.A.M.; Wieringa, B.; Monnens, L.A.H.; van Os
Science 264, 92-95, 1994
A;Title: Requirement of human renal water channel aquaporin-2 for vasopressin-dependent
A;Reference number: A53442; MUID:94188717; PMID:8140421
A;Accession: A53442
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-271 <DEE>
A;Cross-references: UNIPROT:P41181; GB:Z29491; NID:g474058; PIDN:CAA82627.1; PID:g474059
R;Uchida, S.; Sasaki, S.; Fushimi, K.; Marumo, F.
J. Biol. Chem. 269, 23451-23455, 1994
A;Title: Isolation of human aquaporin-CD gene.
A;Reference number: I51877; MUID:95029363; PMID:7524315
A;Accession: I51877
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-186,'C',.188-271 <LIE>
A;Cross-references: GB:S73196; NID:g685000; PIDN:AAB31999.1; PID:g685001
A;Genetics:
A;Gene: GDB:AQP2
A;Cross-references: GDB:141853; OMIM:107777
A;Map position: 12q13-12q13
A;Introns: 120/3; 175/3; 202/3
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: transmembrane protein; water channel

Alignment Scores:
Pred. No.:   271
Score: 597.00
Percent Similarity: 69.37%   Length: 271
Best Local Similarity: 51.35%   Matches: 114
Query Match: 27.37%   Conservative: 40
DB: Mismatches: 64
                                     Indels: 4
                                     Gaps: 2

US-10-723-180-1 (1-1152) x A53442 (1-271)

Qy   136 GCTTTCTGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTTATTTTTGTTCTCCTCAGC 195
Db   8 AlaPheSerArgAlaValPheAlaGluPheLeuAlaThrLeuPhePheGly 27

Qy   196 CTGGGATCACCACCAACTGGGGTGGAACAGAAAAGCCTTACCGGTCGACATGGTTCTC 255
Db   28 LeuGlySerAlaLeuAsnTrp------ProGlnAlaLeuPro---SerValLeuGln 43

Qy   256 ATCTCCCTTTGCTTTGGACTTCAGCATTGCAACCATGGTGCAACATGGTGCAGTGCTTTGGCCATATCAGC 315
Db   44 IleAlaMetAlaPheGlyLeuuGlyIleGlyThrLeuuValGlnAlaLeuuGlyHisIleSer 63

US-10-723-180-1 (1-1152) x JT0750 (1-271)

```
Qy   136 GCTTTCTGGAAAGACTCACAGCGGAATTTCTGGCCATGCTTATTTTGTCTCCTCAGC 195
             ||||||||||  :::||||||||||||||||||||||||||||  |||::||||
Db     8 AlaPheSerArgAlaLeuAlaLeuGluPheLeuAlaThrLeuPheVal PhePheGly 27

Qy   196 CTGGGATCCACCATCAACTCGGGTGGAACAGAAAAGCCTTACCGGTCGACATGGTTCTC 255
             ||||||||  :::  |||          ::::::
Db    28 LeuGlySerAlaLeuGlnTrpAlaSerSer----------ProProSerValLeuGln 43

Qy   256 ACTCCCTTTGCTTTGGACTCAGCATTGCAACCATGGCAGTGCTTTGGCCATATCAGC 315
             ||||||||||||||||||||   ::| |||||  |||::|||
Db    44 IleAlaValAlaAlaPheGlyLeuGlyIleGlyIleLeuValGlnAlaLeuGlyHisValSer 63

Qy   316 GGTGGCCACATCAACCCTGGTCATGGTCTCCGTGAGTGACGTGTGCACCAGGAAGATCAGCATC 375
             ||| ||||||||||||||||||||||     ::  |||||||||||||||  ::|||
Db    64 GlyAlaHisIleAsnProAlaValThrValAlaCysLeuValGlyCysHisValSerPhe 83

Qy   376 GCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCCATCATTGGAGCAGGAATC 435
             ::::::   ||||||||||:::|||   |||||||||||||||||||||
Db    84 LeuArgAlaAlaPheTyrValAlaAlaAlaGlnAlaLeuLeuGlyAlaValAlaAlaIle 103

Qy   436 CTCTATCTGGTCACACCTCCCAGTGGTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGA 495
             ||||| :::  ::  |||||||||   :::       :::
Db   104 LeuHisGluIleThrProValGluIleArgGlyAspLeuAlaValAlaAsnAlaLeuHisAsn 123

Qy   496 AATCTTACCGGTCGCTGGTCATGGTCTCCTGGTTGAGTTGATAATCACATTTCAATTGGTGTTT 555
             ||||||||||||  :::   ||||||||||||||||||||
Db   124 AsnAlaThrAlaGlyGlnAlaValThrValGluLeuPheLeuThrMetGlnLeuValLeu 143

Qy   556 ACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCTCAATAGCTTTAGCA 615
             ||||||||||||||  ||||   :::||   ||
Db   144 CysIlePheAlaSerThrArgGluArgArgGlyAsnLeuGlySerProAlaLeuSer 163

Qy   616 ATTGGATTTTCTGTTGCAATTGCCAGTTATCATTATTTGCAATCAATTATTACTGTGCCAGCATG 675
             ||||||||||||  ||||||||||||||:::|||||||||
Db   164 IleGlyPheSerValThrLeuGlyHisLeuLeuGlyIleTyrPheThrGlyCysSerMet 183

Qy   676 AATCCCGCCGATCCTTTGGACCTGCCAGTTATCATGGGAAATTGGGAAAACCATTGGATA 735
             ||||||||||||  :::::    ||||||||||:::
Db   184 AsnProAlaArgSerLeuAlaProAlaValThrGlyLysPheAspAspHisTrpVal 203

Qy   736 TATTGGGTTGGGCCCATCATAGAGAGCTGTCCTCGGTGGCCTTTATGACTATGTCTTC 795
             :::|||||||||  :::|||||||||
Db   204 PheTrpIleGlyProLeuValGlyAlaIleIleGlySerLeuTyrAsnTyrLeuLeu 223

Qy   796 TGTCCAGATGTTGAA 810
             ||| :::
Db   224 PheProSerAlaLys 228
```

RESULT 10
I51164
chip aquaporin - edible frog
C;Species: Rana esculenta (edible frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51164
R;Abrami, L.; Simon, M.; Rousselet, G.; Berthonaud, V.; Buhler, J.M.; Ripoche, P.
Biochim. Biophys. Acta 1192, 147-151, 1994
A;Title: Sequence and functional expression of an amphibian water channel, FA-CHIP: a ne
A;Reference number: I51164; MUID:94264032; PMID:7515688
A;Accession: I51164
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-272 <ABR>
A;Cross-references: UNIPROT:P50501; GB:L24754; NID:g404778; PIDN:AAC38016.1; PIDN:g404779
C;Genetics:
A;Gene: AQPA
C;Superfamily: lens fiber membrane major intrinsic protein

Alignment Scores:
Pred. No.:   4.14e-41   Length: 272
Score: 589.50   Matches: 122
Percent Similarity: 61.43%   Conservative: 50

```
Qy   316 GGTGGCCACATCAACCCTGGTCATGGTCTCCGTGAGTGACGTGTGCACCAGGAAGATCAGCATC 375
             ||| ||||||||||||||||||||||     ::      :::
Db    64 GlyAlaHisIleAsnProAlaValThrValAlaCysLeuValGlyCysHisValSerVal 83

Qy   376 GCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGCCATCATTGGAGCAGGAATC 435
             ::::::   ||||||||||:::|||   |||||||||||||||||||||
Db    84 LeuArgAlaAlaPheTyrValAlaAlaAlaGlnAlaLeuLeuGlyAlaValAlaAlaLeu 103

Qy   436 CTCTATCTGGTCACACCTCCCAGTGGTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGA 495
             ||||| :::||:::||| |||||||||   :::
Db   104 LeuHisGluIleThrProAlaAspAlaArgGlyAspLeuAlaValAlaAsnAlaLeuSerAsn 123

Qy   496 AATCTTACCGGTCGCTGGTCATGGTCTCCTGGTTGAGTTGATAATCACATTTCAATTGGTGTTT 555
             :::   |||||||||||   ::|||   ||||||||||||
Db   124 SerThrThrAlaGlyGlnAlaValThrValGluLeuPheLeuThrGlnLeuLeuValLeu 143

Qy   556 ACTATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCTCAATAGCTTTAGCA 615
             ||||| :::        ||||:::||  |||::   ||
Db   144 CysIlePheAlaSerThrArgAspArgArgGlyGlyAsnProGlyThrProAlaLeuSer 163

Qy   616 ATTGGATTTTCTGTTGCAATTGGACAATTAATTTCCAATCAATAATACTGGTGCCAGCATG 675
             |||||||||||  :::||   :::|||||||||||
Db   164 IleGlyPheSerValAlaLeuGlyHisLeuLeuGlyIleHisTyrThrGlyCysSerMet 183

Qy   676 AATCCCGCCGATCCTTTATTTGCAATTGGAAATTGGGAAAAACCATTGGATA 735
             |||||||||||  :::   ||||:::||||||
Db   184 AsnProAlaArgSerLeuAlaProAlaValThrGlyGlyLysPheAspAspHisTrpVal 203

Qy   736 TATTGGGTTGGGCCCCATCATAGAGAGCTGTCCTCGGTGGCCTTTATGAGTATGTCTTC 795
             :::||||||||||:::||||   ||||||
Db   204 PheTrpIleGlyProLeuValGlyAlaIleIleLeuLeuGlySerLeuLeuTyrAsnTyrValLeu 223

Qy   796 TGTCCA 801
             |||
Db   224 PhePro 225
```

RESULT 9
JT0750
water channel protein WCH-CD - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: JT0750; S29719
R;Ma, T.; Frigeri, A.; Skach, W.; Verkman, A.S.
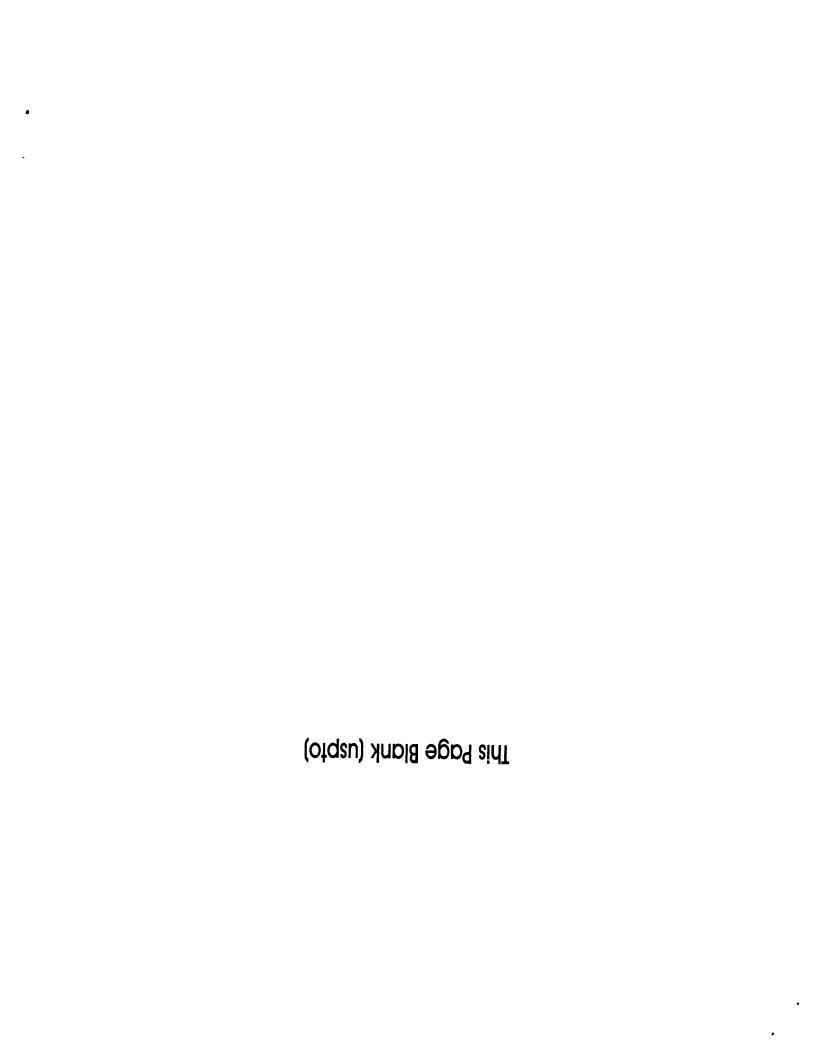Biochem. Biophys. Res. Commun. 197, 654-659, 1993
A;Title: Cloning of a novel rat kidney cDNA homologous to CHIP28 and WCH-CD water channe
A;Reference number: JT0749; MUID:94092144; PMID:7505572
A;Accession: JT0750
A;Molecule type: mRNA
A;Residues: 1-271 <MAT>
A;Experimental source: kidney
R;Fushimi, K.; Uchida, S.; Hara, Y.; Hirata, Y.; Marumo, F.; Sasaki, S.
Nature 361, 549-552, 1993
A;Title: Cloning and expression of apical membrane water channel of rat kidney collectin
A;Reference number: S29719; MUID:93156843; PMID:8429910
A;Accession: S29719
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-271 <FUS>
A;Cross-references: GB:D13906; NID:g286195; PIDN:BAA03006.1; PID:g286196
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: glycoprotein; transmembrane protein
F;124/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:   1.74e-41   Length: 271
Score: 594.00   Matches: 111
Percent Similarity: 66.67%   Conservative: 39
Best Local Similarity: 49.43%   Mismatches: 71
Query Match: 27.24%   Indels: 4
DB: Gaps: 1

```
Best Local Similarity:   43.57%          Mismatches:  77
Query Match:             27.03%          Indels:      31
DB:                          2           Gaps:         5

US-10-723-180-1 (1-1152) x I51164 (1-272)

Qy  133  CAAGCTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTTCTCCTC          192
         :::||||||||||::::|||||:::|||||||||||||||::::|||::::|||||
Db    8  LysAlaPheTrpArgAlaVallleAlaGluPheLeuAlaMetIleLeuPheValPheIle    27

Qy  193  AGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTC----------  243
         ||:::|||||      :::  :::                -------PheProIleGluGluLys  42
Db   28  SerIleGlyAlaAlaLeuGlyPheAsn----------                           42

Qy  244  --------GACATGGTTCTCATCTCCCTTTGCCTTTGGACTC                       276
         :::|||||||       ::|||||||
Db    43  AlaAsnGlnThrValGlyArgSerGlnAlaPheIleValLysValSerLeuAlaPheGlyIle  62

Qy  277  AGCATTGCAACCATGGTGCAGTGCTTTGCCCATATCAGCGGGTGGCCACATCAACCCTGA    336
         ||||||||||||               :::||||||
Db   63  SerIleAlaThrMetAlaGlnIleSerValGlyHisValSerGlyAlaHisLeuAsnProAla  82

Qy  337  GTGACTGTGGCCATGGTGCGCACCAGGAAGATCAGCATTCGCACCAAGTCTGTCTTCTACATC  396
         ||||||              :::  ::|||||||
Db   83  ValThrLeuGlyCysLeuLeuSerCysGlnIleSerIleLeuSerIleLeuLysAlaValMetTyrIle 102

Qy  397  GCAGCCCAGTGCCTGGGGCCATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCC      456
         ||||||                    :::|||||||
Db  103  IleAlaGlnCysLeuGlyLeuGlyAlaValAlaLeuAlaThrAlaIleLeuSerGlyIleThrSerGly 122

Qy  457  AGTGTGGTGGGAGGCCCTGGAGTCCATGGTTCATGGAAATCTTACCGCTGGTCATGGT      516
         :::|||||||
Db  123  LeuGluAsnAsnSerLeuGlyLeuSerProGlyValSerAlaGlyValGlnGly          142

Qy  517  CTCCTGGTTGAGTTGATAATCACATTTCAATTGGTGTTTACTATCTTTGCCAGCTGTGAT    576
         :::|||||||
Db  143  LeuGlyValGluLeuThrPheGlnLeuValLeuCysValValAlaValThrAsp          162

Qy  577  TCCAAACGGACTGATGTCACTGGTGGGGCCATCATTGGATTTCTGTTGCAATT          636
         ||||||
Db  163  ArgArgArgHisAspValSerGlyValProLeuAlaIleGlyLeuSerValAlaLeu       182

Qy  637  GGACATTATTTGCAATCAATAATACTGTGCCAGCATGAATCCCGCCATCCTTTGGA      696
         ||||||
Db  183  GlyHisLeuIleAlaIleLeuAspTyrTyrGlyCysGlyMetAsnProAlaArgsgSerPheGly 202

Qy  697  CCTGCAGTTATCATGGGAAAATTGGGAAAACCATTGGGATTATATGGGTTGGGGCCCCATCATA  756
         :::||||||
Db  203  SerAlaValLeuThrLysAsnPheThrTyrHisTrpIlePheTrpValGlyProMetIle        222

Qy  757  GGAGCTGTCCTCGGTGGCCTTATGGTATGTCTTCGTGTCGCCAGATGTTGAATTCAAA    816
         ||:||
Db  223  GlyGlyAlaAlaAlaAlaAlaIleIleTyrAspPheIleLeuAlaPro----------     237

Qy  817  CGTCGTTTAAAGAAGCCTTCAGCAAGCTGCCCAGCAAACAAAGGAAGCTACATGGAG    876
         :::||
Db  238  --ArgThrSerAspLeuThrAspArgMetLysValTrpThrAsnGlyLysValGluGlu 256

Qy  877  GTCGAG---------GACAACAGGAGTCAGGTAGAGACCGGATGACCCTGATTCTAAAACCT  927
         :::|||||
Db  257  TyrGluLeuAspGlyAspAsnThrArgValGlu----------MetLysProLys         271
```

RESULT 11

I64818

water-channel aquaporin 2 - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 13-Aug-1999
C;Accession: I64818
R;van Lieburg, A.F.; Verdijk, M.A.; Knoers, V.V.; van Essen, A.J.; Proesmans, W.; Mallma
Am. J. Hum. Genet. 55, 648-652, 1994
A;Title: Patients with autosomal nephrogenic diabetes insipidus homozygous for mutations
A;Reference number: I51877; MUID:95029363; PMID:7524315
A;Accession: I64818

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-271 <RBS>
A;Cross-references: GB:S73197; NID:g684998; PIDN:AAB31998.1; PID:g684999
C;Genetics:
C;Gene: GDB:AQP2
A;Cross-references: GDB:141853; OMIM:107777
A;Map position: 12q13-12q13
C;Superfamily: lens fiber membrane major intrinsic protein

```
Alignment Scores:
Pred. No.:                                Length:        271
Score:                   4.56e-41         Matches:       113
Percent Similarity:      589.00           Conservative:   40
Best Local Similarity:   68.92%           Mismatches:     65
Query Match:             50.90%           Indels:          4
DB:                      27.01%           Gaps:            2
                             2

US-10-723-180-1 (1-1152) x I64818 (1-271)

Qy  136  GCTTTCTGGAAAGCAGTCACAGCCGGAATTTCTGGCCATGCTTATTTTGTTCTCCTCAGC      195
         ||:::|||||||          :::  :::|||||||          :::
Db    8  AlaPheSerArgAlaValPheAlaGluPheAlaGluLeuAlaThrLeuLeuPheValPhePheGly 27

Qy  196  CTGGGATCCACCATCAACTGGGGTGGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTC      255
         ||||||                    :::               ProGlnAlaGlyAsnTrp----
Db   28  LeuGlySerAlaLeuAsnTrp----------          ---SerValLeuGln          43

Qy  256  ATCTCCCTTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGCCCATATCAGC    315
         ||:::              ::|||||||
Db   44  IleAlaMetAlaPheGlyLeuGlyIleIleGlyThrLeuGlnAlaLeuGlyHisIleSer     63

Qy  316  GGGTGGCCACATCAACCCTGAGTGACTGTGGCCATGGTGCGCACCAGGAAGATCAGCATC    375
         |||||                    :::|||||||
Db   64  ArgAlaHisIleAsnProAlaValThrLeuGlyCysLeuValGlyCysHisValSerVal     83

Qy  376  GCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGCCATCATTGGAGCAGGAATC    435
         ||||||
Db   84  LeuArgAlaAlaPheTyrValAlaAlaGlnLeuLeuGlyLeuGlyAlaValAlaAlaLeu    103

Qy  436  CTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCCTGGAGTTGAGTTGATAATCACA    495
         ||||||              :::
Db  104  LeuHisGluIleThrProAlaAspIleArgGlyAlaValAlaAlaValAlaLeuSerAsn   123

Qy  496  AATCTTACCGCTGGTCATGGTCCTGGTTGAGTTGAGTGATAATCACATTTCAATTGGTGTT    555
         ::: |||||||
Db  124  SerThrThrAlaGlyValGlnAlaValAlaValThrArgArgArgGlyGlyAsnProAlaArg  143

Qy  556  ACTATCTTTGCCAGCTGTGATTCCAACCGACTGATGTCACTGGTCACTGGTCAATAGCTTTAGCA  615
         :::|||
Db  144  CysIlePheAlaSerThrArgAspGluArgArgGlyValGlyAsnProGlyThrProAlaLeuSer 163

Qy  616  ATTGGATTTTCGTTGCAATTGGACATTATTTTGCAATCAATAATACTGTGCCAGCATG    675
         ||||||
Db  164  IleGlyPheSerValAlaLeuGlyHisLeuLeuGlyIleHisTyrThrGlyCysSerMet   183

Qy  676  AATCCCGCCCGATCCTTTGGACTTGTCAGTTATCATGGGAAAATTGGGAAAACCATTGGATA    735
         ||||||
Db  184  AsnProAlaArgSerLeuAlaProAlaValAlaThrGlyLysPheAspAspHisTrpIleVal   203

Qy  736  TATTGGGTTGGGCCCCATCATAGGAGCTGTCCTCGGTGGCCTTATGGTATGTCTTC       795
         ||:||
Db  204  PheTrpIleGlyProLeuValGlyAlaIleLeuLeuGlySerLeuLeuTyrAsnTyrValLeu  223

Qy  796  TGTCCA      801
         ||:||
Db  224  PhePro      225
```

RESULT 12

B44499

major intrinsic protein homolog DER2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: B44499
R;Lanahan, A.; Williams, J.B.; Sanders, L.K.; Nathans, D.
Mol. Cell. Biol. 12, 3919-3929, 1992
A;Title: Growth factor-induced delayed early response genes.
A;Reference number: A44499; MUID:92375060; PMID:1508193
A;Accession: B44499
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-269 <LAN>
A;Cross-references: UNIPROT:Q02013; GB:L02914; NID:g193642; PIDN:AAB53928.1; PID:g193643
A;Experimental source: BALB/c 3T3 cells
A;Note: sequence extracted from NCBI backbone (NCBIP:11645)
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: transmembrane protein

Alignment Scores:
Pred. No.:   269
Score: 9.86e-40   Length: 269
Percent Similarity: 573.00   Matches: 123
Best Local Similarity: 63.43%   Conservative: 47
Query Match: 45.90%   Mismatches: 84
DB: 26.27%   Indels: 14
                          2   Gaps: 6

US-10-723-180-1 (1-1152) x B44499 (1-269)

Qy   139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTTGTTCCTCCAGCCTG 198
           ::|||| :| |:|||||||||||||| ::||||| ::||||||
Db   10   PheTrpArgAlaValValAlaGluPheLeuAlaMetThrLeuPheValPheIleSerIle 29

Qy   199 GGATCCACCATCAACTGGGT--------GGAACAGAAAAGCCTTTACCGGTCGACATG 249
           |||||| ::   ::| |||
Db   30   GlySerAlaLeuGlyPheAsnTyrProLeuGluArgAsnGlnThrLeuValGlnAspAsn 49

Qy   250 GTTCTCATCTCCCTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGCCAT 309
           || ::||||||| ||||||| ||| ||| ||| ||||||||||
Db   50   ValLysValSerLeuAlaPheGlyLeuSerIleLeuAlaThrLeuAlaGlnSerValGlyHis 69

Qy   310 ATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGCACCAGAAGATC 369
           ||||||||||||||||||||| ::||| |||::: || || :::||||
Db   70   IleSerGlyAlaHisLeuAsnProAlaValThrLeuValLeuLeuSerCysGlnIle 89

Qy   370 AGCTACTCGCCAAGTCGTCTTCTCTACACTGCCGAGCGCCCAGTGCCTGGGGGCCATCAATTGGAGCA 429
           |||||| :||||||||||| :::|||||||||||||||||::|||
Db   90   SerIleLeuArgAlaValMetTyrIleIleLeuAlaGlnCysValGlyValAlaIleValAlaThr 109

Qy   430 GGAATCCCTCTATCTGGTCACACCTCCCAGTGTGGTGGGAGGCCTGGAGTCACC---ATG 486
           ||||| |||| :::| ||||||||||| |||
Db   110 AlaIleLeuSerGlyIleThrSerSerLeuValAspAsnSerLeuArgAsnGlyAspLeu 129

Qy   487 GTTCATGGAAATCTTACCGGCTGGTCATGGTCTCCTGGTTGAGTTGATAATCACATTTCAA 546
           || ::|| ::|||||:::||||||||||||| ::|||::|||
Db   130 AlaHisGly--ValAsnSerGlyGlnGlyLeuTyrIleIleLeuGlyIleIleGlyThrLeuGln 148

Qy   547 TTGGTGTTTACTATCTTTGACCTGTGATCCAAACGACTGATGTCACTGGCTCAATA 606
           |||| ::|:::|||||||||::|||
Db   149 LeuValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlyLeuSerAla 168

Qy   607 GCTTTAGCAATTGGATTTTCTGTTGCAATTGGACATTTATTTGCAATCAATTACTGGT 666
           ||||||| ::::||||||||||::|||| |||:::|||||||:::
Db   169 ProLeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuAlaIleAsnAspTyrThrGly 188

Qy   667 GCCAGCATGAAATCCCGCCGCCATCCTTTGACCTGCAGTTATCATGGGAAAAAC 726
           |||||||| |||||||||||::||||||||| ||||||||
Db   189 CysGlyIleAsnProAlaArgSerPheGlySerAlaValAlaLeuThrArgAsnPheSerAsn 208

Qy   727 CATTGGGATATATTGGGTTGGGCCCCATCATAGGAGCTGTCTCCGCTGGTGGCCTTTATGAG 786
           ||||||| ::::||||| ||||||||||||||:::|||::||||:::
Db   209 HisTrpIlePheTrpValGlyProPheIleGlyAlaAlaLeuAlaValLeuIleTyrAsp 228

Qy   787 TATGTCTTCTGTCCA---GATGTTGAATTCAAACTCGTTTTAAAGAAGACCCTTCAGCAAA 843
           ::|:::|||| |||||| ||| ||||:::|||
Db   229 PheIleLeuAlaProArgSerArgAspPheThrArgAspArgMetLysValTrpThrSerGly 248

           ::::   ::::| ||||::: :::|||||||||||||||||||||||
Db   249 GlnValGluGlyTyrAsp-------LeuAspAlaAspAspIleAsnSerArgValGlu 265

Qy   904 ACGGATGACCTGATTCTAAAACCT 927
                                  ::|||
Db   266 ------------MetLysPro 268

RESULT 13
JC1320
water channel protein CHIP28 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: JC1320; JT0749; S37639; A44395
R;Deen, P.M.T.; Dempster, J.A.; Wieringa, B.; Van Os, C.H.
Biochem. Biophys. Res. Commun. 188, 1267-1273, 1992
A;Title: Isolation of a cDNA for rat CHIP28 water channel: High mRNA expression in kidney
A;Reference number: JC1320; MUID:93075224; PMID:1280133
A;Accession: JC1320
A;Molecule type: mRNA
A;Residues: 1-269 <DEE>
A;Cross-references: UNIPROT:P29975; GB:X67948; NID:g55941; PIDN:CAA48134.1; P
A;Experimental source: kidney
R;Ma, T.; Frigeri, A.; Skach, W.; Verkman, A.S.
Biochem. Biophys. Res. Commun. 197, 654-659, 1993
A;Title: Cloning of a novel rat kidney cDNA homologous to CHIP28 and WCH-CD water channe
A;Reference number: JT0749; MUID:94092144; PMID:7505572
A;Accession: JT0749
A;Molecule type: mRNA
A;Residues: 1-4,'F',6-121,'SN',124-269 <MAT>
A;Experimental source: kidney
R;Shanahan, C.M.; Weisberg, P.L.; Metcalfe, J.C.
Circ. Res. 73, 193-204, 1993
A;Title: Isolation of gene markers of differentiated and proliferating vascular smooth m
A;Reference number: S37637; MUID:93284726; PMID:8508530
A;Accession: S37639
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2,'T',4-158,'Q',160-269 <SHA>
A;Cross-references: EMBL:X71069; NID:g313803; PIDN:CAA50395.1; PID:g313804
R;Zhang, R.; Skach, W.; Hasegawa, H.; van Hoek, A.N.; Verkman, A.S.
J. Cell Biol. 120, 359-369, 1993
A;Title: Cloning, functional analysis and cell localization of a kidney proximal tubule
A;Reference number: A44395; MUID:93132068; PMID:8421053
A;Accession: A44395
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-4,'F',6-269 <ZHA>
A;Cross-references: GB:X70257; GB:S52660; NID:g312907; PIDN:CAA49761.1; PID:g312924
A;Experimental source: renal cortex
A;Note: sequence extracted from NCBI backbone (NCBIP:122674)
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: glycoprotein; transmembrane protein
F;42,205/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:   269
Score: 1.2e-39   Length: 269
Percent Similarity: 572.00   Matches: 120
Best Local Similarity: 63.67%   Conservative: 50
Query Match: 64.94%   Mismatches: 85
DB: 26.23%   Indels: 12
                          2   Gaps: 4

US-10-723-180-1 (1-1152) x JC1320 (1-269)

Qy   139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGTTTATTTTTGTTCCTCCAGCCTG 198
           ::|||| :| |:|||||||||||||| ::||||| ::||||||
Db   10   PheTrpArgAlaValValAlaGluPheLeuAlaMetThrLeuPheValPheIleSerIle 29

Qy   199 GGATCCACCATCAACTGGGT--------GGAACAGAAAAGCCTTTACCGGTCGACATG 249
           |||||| ::   ::| |||
Db   30   GlySerAlaLeuGlyPheAsnTyrProLeuGluArgAsnGlnThrLeuValGlnAspAsn 49

DB:

US-10-723-180-1 (1-1152) x JC2348 (1-271)

```
                    2                           Gaps:                      5

Qy   139 TTCTGGAAAGCAGTCACAGCGGGAATTCTGGCCATGCTATTTTGTCTCCTCAGCCTG 198
             |||||.:|||||  ||||||:|||||||||||||  ::: :::||||.||||||
Db    10 PheTrpArgAlaValValAlaGluPheLeuAlaMetIleLeuPheIleSerIle 29

Qy   199 GGATCCACCATCAACTGGGGTGGAACAGAAAAGCTTACCG------------------ 240
             ::                          ::::
Db    30 GlySerAlaLeu--------GlyPheHisTyrProIleLysSerGlnThrThrGly 46

Qy   241 ------GTGCCAGCTGTTCATCCTCCCTTGCTTGGACTCAGCTCAGCATTGCAACCATGGTG 294
                   |||||||||  :::
Db    47 AlaValGlnAspAsnValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAla 66

Qy   295 CAGTGCTTGGCCATATCACGGTGGCCACATCAACCCTGACGTGGCCATGGTG 354
             |||||||||||||||  :: :::||||
Db    67 GlnSerValGlyHisIleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeu 86

Qy   355 TGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACACGCAGCCCAGTGCTGTGGGG 414
             |||||||||
Db    87 LeuSerCysGlnIleSerValLeuLeuArgAlaIleLeuMetTyrIleIleGlnCysValGly 106

Qy   415 GCCATCATTGGAGCAGGAATCCTCTATCGGCAATCCCCCAGTGTGTGGGAGGCCTG 474
             |||||||
Db   107 AlaIleValAlaThrAlaLeuSerGlyIleThrSerSerLeuProAspAsnSerLeu 126

Qy   475 GGAGTCACCATGGTTCATGGAAATCTTACCGCTCGTCCTCGGTTGAGTTGATA 534
             |||||||  :::                        :::
Db   127 GlyLeuAsnAlaLeuAlaProGlyValAsnSerGlyAlyValAsnSerGlyLeuGlyLeuIleGluIleIle 146

Qy   535 ATCACATTTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTC 594
             |||  |||  :::                          :::    ||||||
Db   147 GlyThrLeuGlnLeuValLeuCysGlyValLeuLeuAlaThrThrAspArgAspLeu 166

Qy   595 ACTGGCTCAATAGCTTAGCAATTGGATTTCTTGCAATTGGACATTTATTATTGCAATC 654
             |||||||                        ||||||||  |||||||
Db   167 GlyGlySerGlyProLeuAlaIleGlyPheSerValAlaLeuGlyHisLeuLeuAlaIle 186

Qy   655 AATTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGA 714
             ::||||||||                          :::||||||
Db   187 AspTyrThrGlyCysGlyIleAsnProAlaArgSerPheSerValIleIleThrHis 206

Qy   715 AATTGGGAAAACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTGCTGGT 774
             ::|||||||                          :::|||||
Db   207 AsnPheGlnAspHisTrpIlePheTrpValGlyProPheIleGlyProPheIleGlyAlaLeuAlaLeuAlaVal 226

Qy   775 GGCCTTTATGGATGAGTATGTCTTCTGTGTCCAGATGTGAATTCAAACGTCGTTTAAAGAAGCC 834
             ::||||||.:                          :::
Db   227 LeuIleTyrAspPheIleLeuLeuAlaPro-----------------ArgSerAspLeu 240

Qy   835 TTCAGCAAAGTGCGCCAGCAAACAAAAGGAAGCTAC------------ATGGAGGTGGAG 882
             ::                          ||| |||||||
Db   241 ThrAspArgValLysValTrpThrSerGlyGlnValGluTyrAspLeuAspAlaAsp 260

Qy   883 GACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCT 927
             :: :||||||
Db   261 AspIleAsnSerArgValGlu---------MetLysPro 270
```

RESULT 15

IS2366

uterine water channel - human

C;Species: Homo sapiens (man)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999

C;Accession: IS2366

R;Li, X.; Yu, H.; Koide, S.S.

Biochem. Mol. Biol. Int. 32, 371-377, 1994

A;Title: The water channel gene in human uterus.

A;Reference number: IS2366; MUID:94290349; PMID:7517253

A;Accession: IS2366

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

```
Qy   250 GTTCTCATCTCCCTTGCTTGGACTCAGCTCAGCATTGCAACCATGGTGCAGTGCTTGGCCAT 309
             ||||  ::::||||  ||||||||||||||  ||||||
Db    50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69

Qy   310 ATCAGCGGTGGCCACATCAACCCTGACGTGGCCATGGTGTGCACCAGGAAGATC 369
             |||||||||  :: :::  ||||||  :: :::||||
Db    70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuLeuSerCysGlnIle 89

Qy   370 AGCATCGCCAAGTCTGTCTTCTACACGCAGCCCAGTGCTGTGGGGCCATCATTGGAGCA 429
             |||||||||||||  :::                 :::||||
Db    90 SerIleLeuArgAlaIleMetTyrIleIleGlnCysValGlyAlaIleValAlaThrAlaSer 109

Qy   430 GGAATCCTCTATCGGTCACACCTCCCAGTGTGTGGGGAGGCCTGGGAGTCACCATGGTT 489
             |||  ::: :::
Db   110 AlaIleLeuLeuSerGlyIleThrSerSerLeuProAspAsnSerLeuGlyArgAspLeu 129

Qy   490 CATGGAAATCTTACCGGTCGTCCTCGGTTGAGTTGATAATCACATTTCAATTG 549
             ||||||||                          :::||||  :: :::
Db   130 AlaArgGlyValAsnSerGlyValGlnGlyLeuGlyLeuIleGluIleIleGlyThrLeuGlnLeu 149

Qy   550 GTGTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGATGTCACTGGCTCAATAGCT 609
             ::: :::                          :::||||  ||||||
Db   150 ValLeuCysGlyValLeuLeuAlaThrThrAspArgArgAspLeuGlyGlySerAlaPro 169

Qy   610 TTAGCAATTGGATTTCTTGCAATTGGACATTTATTATTGCAATCAATTATACTGGTGCC 669
             |||||||                          ||||||  :::  ||||||
Db   170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189

Qy   670 AGCATGAATCCCGCCGATCCTTTGGACCTGTATCATGGGAAATTGGGAAAACCAT 729
             |||||||||  :: :::                          ||||||
Db   190 GlyIleAsnProAlaArgSerPheGlySerAlaValAlaThrArgAsnSerAsnHis 209

Qy   730 TGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTGCGCTTATGAGTAT 789
             ||||||  :: :::                          :::||||||
Db   210 TrpIlePheTrpValGlyProPheIleGlyProPheIleGlySerAlaLeuAlaValLeuIleTyrAspPhe 229

Qy   790 GTCTTCTGTCCA---GATGTTGAATTCAAACGTCGTTTAAAGAAGCCTTCAGCAAAGCT 846
             :::                          ||| |||||||
Db   230 IleLeuLeuAlaProSerArgSerAspPheThrAspArgMetLysValTrpThrSerGlyGln 249

Qy   847 GCCCAGCAAACAAAAGGAAGCTACATGGAGGTGGAGGACAACAGGAGTCAGGTAGAGACG 906
             |||||||  :: :::                          ||||||
Db   250 ValGluTyrAsp---------LeuAspAlaAspAspIleAsnSerArgValGlu--- 265

Qy   907 GATGACCTGATTCTAAAACCT 927
             :: :||||||
Db   266 ---------MetLysPro 268
```

RESULT 14

JC2348

water channel protein CHIP29 - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C;Accession: JC2348

R;Patil, R.V.; Yang, X.; Saito, I.; Coca-Prados, M.; Wax, M.B.

Biochem. Biophys. Res. Commun. 204, 861-866, 1994

A;Title: Cloning of a novel cDNA homologous to CHIP28 water channel from ocular ciliary

A;Reference number: JC2348; MUID:95071406; PMID:7526655

A;Accession: JC2348

A;Molecule type: mRNA

A;Residues: 1-271 <PAT>

A;Cross-references: UNIPROT:P47865; GB:S74759; NID:g786448; PIDN:AAB32365.1; PID:g7864500

A;Experimental source: ciliary epithelium

C;Superfamily: lens fiber membrane major intrinsic protein

C;Keywords: glycoprotein

F;42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	2.13e-39	Length:	271
Score:	569.00	Matches:	118
Percent Similarity:	61.82%	Conservative:	52
Best Local Similarity:	42.91%	Mismatches:	79
Query Match:	26.09%	Indels:	26

A;Residues: 1-269 <RES>
A;Cross-references: GB:S73482; NID:g688357; PIDN:AAB31193.1; PID:g688358
C;Superfamily: lens fiber membrane major intrinsic protein

Alignment Scores:
Pred. No.:   Length: 269
Score: 6.74e-39   Matches: 118
Percent Similarity: 61.11%   Conservative: 47
Best Local Similarity: 43.70%   Mismatches: 87
Query Match: 25.81%   Indels: 18
DB: Gaps: 4

US-10-723-180-1 (1-1152) x I52366 (1-269)

Qy 139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGCCATGCTTATTTTGTTCTCCTCAGCCTG 198
        |||||||:::||||| |||||||||||| |||| :::|||||| :||||
Db   10 PheTrpArgAlaValAlaGluPheLeuAlaThrThrLeuPheValPheIleSerIle 29

Qy 199 GGATCCACCATCAACTGGGGT--------GGAACAGAAAAGCCTTTACCGGTCGACATG 249
        ||| |::| ||| ::: |||||||||||
Db   30 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrValValGlnAspAsn 49

Qy 250 GTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTCAGCTGCTTTGGCCAT 309
        ||| :::||||||||| |||| |||::: ||||||||||||| |||||
Db   50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69

Qy 310 ATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGCACCAGGAAGATC 369
        ||||||:||:||:||||| ||| |||||:::   ::: :::||||||||||
Db   70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuLeuSerCysGlnIle 89

Qy 370 AGCATCGCCCAAGTCTGTCTTCTACATCGCAGCCCAGTGCCTGGGGGGCCATCATTGGAGCA 429
        ||||| ::::::: |||||| |||| ||||||:::||||||||||||:||||||:::
Db   90 SerIlePheArgAlaLeuMetTyrIleIleIleAlaGlnCysValGlyAlaIleValAlaThr 109

Qy 430 GGAATCCTCTATCTGCTCACACCTCCCAGTGTGGTGGGAGGCCTGGGAGTCACCATGGTT 489
        ||| ||||||||| :::| |||||||||||||||||
Db 110 AlaIleLeuSerGlyIleThrSerLeuThrProProSerValValGlyArgAsnAspLeu 129

Qy 490 CATGAAATCTTACCGCTGCTGTCATCGCTGCTGGTTGAGTGATAATCACATTTCAATTG 549
        ||| |||||||| ::::|| :::|||||||||:: ||| ||||||||
Db 130 AlaAspGlyValAsnSerGlyLysGlnLeuGlyLysLeuLeuGluLeuIleGlyThrLeuGlnLeu 149

Qy 550 GTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGACTGATGTCACTGGCTCAATAGCT 609
        ||| ::: |||||| ||||| :::|| ||||| ||||||:::|||||||
Db 150 ValLeuCysValLeuAlaThrThrAspArgArgArgArgArgAspLeuGlySerAlaPro 169

Qy 610 TTAGCAATTGGATTTCTGTTGCAATTGACACATTTATTTGCAATCAATTATACTGGTGCC 669
        |||||||||||||||| |||||||::: ||||||||:::|||||:::||||||
Db 170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189

Qy 670 AGCATGAATCCGCCCGATCCTTTGGACCTGCAGTTATCATGGGAAAATTGGGAAAACCAT 729
        |||||||||||||||||| :::::: ||||||| |||||| ||| |||||:::
Db 190 GlyIleLeuAsnProAlaArgSerPheGlySerAlaValIleIleThrArgAsnPheSerAsnHis 209

Qy 730 TGGATATATTGGGTTGGGCCCATCATAGAGCTGTCCTCCGGTGGCCTTTATGAGTAT 789
        |||||||||| ::::||||| |||||||||:::||||| :::::::
Db 210 TrpIlePheTyrValGlyProPheIleGlyValAlaLeuAlaValLeuIleTyrAspPhe 229

Qy 790 GTCTTCTGTCCAGATGTTGAATTCAAAACGTCGTTTTAAAGAAGCCTTCAGCAAAGCTGCC 849
        ::: :::|||::: ||| ||| :::
Db 230 --------ArgSerAspLeuThrAspArgValLys 243

Qy 850 CAGCAACAAAACAGGAAGCTAC---------ATGGAGGTGGAGGAACAACAGGAGTCAG 897
        ||| || :::::: ::|||| |||:::||||||||||
Db 244 ValTrpThrSerGlyGlnValGluGlyTyrAspLeuAspAlaAspIleAsnSerArg 263

Qy 898 GTAGAGACGGATGACCTGATTCTAAAACCT 927
        ::|||||||
Db 264 ValGlu--------MetLysPro 268

Search completed: September 8, 2005, 12:15:27
Job time : 56.5 secs

This Page Blank (uspto)

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 8, 2005, 11:55:12 ; Search time 229 seconds
                                       (without alignments)
                              5152.098 Million cell updates/sec

Title:          US-10-723-180-1
Perfect score:  2181
Sequence:       1 ggggcaggcaatgagagctg........attactcagtctaaacaata 1152
Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:       3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
```

Command line parameters:
```
-MODEL=frame+_n2p.model -DBV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10723180/runat_08092005_123256_29239/app_query.fasta_1.1351
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10723180 @CGN_1_1.248 @runat_08092005_123256_29239 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      UniProt 03:*
            1: uniprot_sprot:*
            2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1685	77.3	323	1	AQP4_HUMAN	P55087 homo sapien
2	1628	74.6	323	2	AQP4_BOVIN	O77750 bos taurus
3	1623	74.4	323	2	Q86654	Q86654 ovis aries
4	1596	73.2	324	1	AQP4_DIPME	Q923J4 dipodomys m
5	1595	73.1	323	1	AQP4_RAT	P47863 rattus norv
6	1567	71.8	323	1	AQP4_MOUSE	P55088 mus musculu
7	1505	69.0	301	2	Q6XVT6	Q6xvt6 ovis aries
8	1466	67.2	335	2	Q8UVB8	Q8uvb8 coturnix co
9	1449	66.4	335	2	Q65YQ3	Q65yq3 gallus gall
10	1103	50.6	224	1	Q8BR89	Q8br89 mus musculu
11	1080.5	49.5	218	1	Q6L7A0	Q6l7a0 homo sapien
12	1037.5	47.6	320	2	Q6AZD2	Q6azd2 brachydanio
13	884	40.5	178	1	O46426	O46426 oryctolagus
14	707	32.4	149	1	Q8K4M1	Q8k4m1 mus musculu
15	656.5	30.1	249	1	AQP_AEDAE	Q9nhw7 aedes aegyp
16	625	28.7	244	2	Q7PWV1	Q7pwv1 anopheles g
17	621	28.5	274	2	Q6T629	Q6t629 coturnix co
18	620	28.4	271	1	Q9YI68	Q9yi68 bufo marinu
19	617	28.3	265	1	AQP5_HUMAN	P55064 homo sapien
20	615	28.2	299	2	Q6AYU6	Q6ayu6 rattus norv
21	614	28.2	265	1	AQP5_MOUSE	Q9wty4 mus musculu
22	614	28.2	265	1	AQP5_RAT	P47864 rattus norv
23	614	28.2	271	1	Q8JIR3	Q8jir3 hyla japoni
24	612	28.1	263	1	MIP_HUMAN	P30301 homo sapien
25	606	27.8	263	1	MIP_BOVIN	P06624 bos taurus
26	605	27.7	271	1	AQP2_SHEEP	O62735 ovis aries
27	605	27.7	273	2	Q8AWH9	Q8awh9 xenopus lae
28	604	27.7	263	2	Q6J8I9	Q6j8i9 ovis aries
29	604	27.7	271	1	AQP2_MOUSE	P56402 mus musculu
30	603.5	27.7	270	2	Q66LN3	Q66ln3 coturnix co
31	603.5	27.7	274	2	Q6IP27	Q6ip27 xenopus lae
32	602.5	27.6	268	2	Q9YI69	Q9yi69 bufo marinu
33	602	27.6	263	2	MIP_RANPI	Q06019 rana pipien
34	600	27.5	261	1	MIP_RAT	P09011 rattus norv
35	599	27.5	262	2	Q8QFW8	Q8qfw8 gallus gall
36	598.5	27.4	274	2	Q6DJ01	Q6dj01 xenopus tro
37	597.5	27.4	263	1	MIP_MOUSE	P51180 mus musculu
38	597	27.4	271	1	AQP2_HUMAN	P41181 homo sapien
39	596.5	27.3	263	2	Q640N8	Q640n8 mus musculu
40	596.5	27.3	268	2	Q76E78	Q76e78 hyla japoni
41	595	27.3	271	2	Q8JIR4	Q8jir4 hyla japoni
42	595	27.3	274	2	Q9YI67	Q9yi67 bufo marinu
43	594	27.2	271	1	AQP2_RAT	P34080 rattus norv
44	594	27.2	271	2	Q6FGT3	Q6fgt3 homo sapien
45	589.5	27.0	272	1	AQPA_RANES	P50501 rana escule

ALIGNMENTS

```
RESULT 1
AQP4_HUMAN
ID   AQP4_HUMAN     STANDARD;      PRT;   323 AA.
AC   P55087; P78564;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Aquaporin 4 (WCH4) (Mercurial-insensitive water channel) (MIWC).
GN   Name=AQP4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=96176324; PubMed=8601457; DOI=10.1016/0014-5793(96)00092-0;
RA   Misaka T., Abe K., Iwabuchi K., Kusakabe Y., Ichinose M., Miki K.,
RA   Emori Y., Arai S.;
RT   "A water channel closely related to rat brain aquaporin 4 is expressed
RT   in acid- and pepsinogen-secretory cells of human stomach.";
RL   FEBS Lett. 381:208-212(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96032721; PubMed=7559426; DOI=10.1074/jbc.270.39.22907;
RA   Yang B., Ma T., Verkman A.S.;
RT   "cDNA cloning, gene organization, and chromosomal localization of a
RT   human mercurial insensitive water channel. Evidence for distinct
RT   transcriptional units.";
RL   J. Biol. Chem. 270:22907-22913(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97008105; PubMed=8855281; DOI=10.1073/pnas.93.20.10908;
RA   Lu M., Lee M.D., Smith B.L., Jung J.S., Agre P., Verdijk M.A.J.,
RA   Merkx G., Rijss J.P.L., Deen P.M.T.;
RT   "The human AQP4 gene: definition of the locus encoding two water
RT   channel polypeptides in brain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:10908-10912(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC       regulates body water balance and mediates water flow within the
CC       central nervous system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=2;
CC         IsoId=P55087-1; Sequence=Displayed;
CC       Name=1;
CC         IsoId=P55087-2; Sequence=VSP_003232;
CC   -!- TISSUE SPECIFICITY: Brain - muscle >> heart, kidney, lung, and
CC       trachea.
CC   -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC       three membrane-spanning domains and a pore-forming loop with the
CC       signature motif Asn-Pro-Ala (NPA).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U34846; AAC52112.1; ALT_INIT.
DR   EMBL; U34845; AAC50284.1; ALT_INIT.
DR   EMBL; D63412; BAA09715.1; -.
DR   EMBL; U63622; AAB26957.1; -.
DR   EMBL; U63623; AAB26958.1; -.
DR   EMBL; BC022286; AAH22286.1; -.
DR   PIR; I39178; I39178.
DR   HSSP; P29972; 1FQY.
DR   Genew; HGNC:637; AQP4.
DR   H-InvDB; HIX0014377; -.
DR   MIM; 600308; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005372; F:water transporter activity; TAS.
DR   GO; GO:0007588; P:excretion; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0006810; P:transport; TAS.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW   Transport.
FT   DOMAIN        1     36       Cytoplasmic (Potential).
FT   TRANSMEM     37     57       Potential.
```

```
FT   DOMAIN       58     64       Extracellular (Potential).
FT   TRANSMEM     65     85       Potential.
FT   DOMAIN       86    115       Cytoplasmic (Potential).
FT   TRANSMEM    116    136       Potential.
FT   DOMAIN      137    155       Extracellular (Potential).
FT   TRANSMEM    156    176       Potential.
FT   DOMAIN      177    184       Cytoplasmic (Potential).
FT   TRANSMEM    185    205       Potential.
FT   DOMAIN      206    231       Extracellular (Potential).
FT   TRANSMEM    232    252       Potential.
FT   DOMAIN      253    323       Cytoplasmic (Potential).
FT   SITE         97     99       NPA 1.
FT   SITE        213    215       NPA 2.
FT   MOD_RES     285    285       Phosphoserine (By similarity).
FT   CARBOHYD    153    153       N-linked (GlcNAc...) (Potential).
FT   CARBOHYD    206    206       N-linked (GlcNAc...) (Potential).
FT   VARSPLIC      1     22       Missing (in isoform 1).
FT                                /FTId=VSP_003232.
FT   CONFLICT    246    246       G -> A (in Ref. 1).
FT   CONFLICT    287    288       VE -> AK (in Ref. 1).
FT   CONFLICT    296    296       P -> L (in Ref. 1).
SQ   SEQUENCE    323 AA;  34829 MW;  1A160CF0DC11052 CRC64;

Alignment Scores:
Pred. No.:                                 Length:              323
Score:                      1.41e-119      Matches:             323
Percent Similarity:         1685.00        Conservative:          0
Best Local Similarity:      100.00%        Mismatches:            0
Query Match:                77.26%         Indels:                0
DB:                         1              Gaps:                  0

US-10-723-180-1 (1-1152) x AQP4_HUMAN (1-323)

Qy     40 ATGAGTGACGACCCACAGCAAGGCGGTGGGTAAGTGTGGACCTTGTGTACCAGAGAG 99
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetSerAspArgProThrAlaArgArgTrpLysGlyProLeuCysThrArgGlu 20

Qy    100 AACATCATGGTGGCTTCAAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCG 159
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40

Qy    160 GAATTCTGCCATGCTATTTTGTTCTCCTCCAGCTGGGATCCACCATCAACTGGGGT 219
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 GluPheLeuAlaMetLeuLeuPheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60

Qy    220 GGAACAGAAAAGCCTTACCGGTGCGACATGGTTCTCATCCCCCTTGCTTGGACTCAGC 279
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuGlyLeuSer 80

Qy    280 ATTGCAACCATGGTGCAGTGCTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG 339
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100

Qy    340 ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATGCCCAAGTCTGTCTTCTACATCGCA 399
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 ThrValAlaMetValCysThrArgLysIleSerMetProLysSerValPheTyrIleAla 120

Qy    400 GCCCAGTGCTCGGGGGCCATCATTGGAGCAGGAGGAATCCTCTATCTGGTCACACCTCCAGT 459
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyGlyIleLeuTyrLeuValThrProProSer 140

Qy    460 GTGGTGGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC 519
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ValValGlyGlyValGlyLeuValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160

Qy    520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTTACTATCTTGCCAGCTGTGATTCC 579
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIleSerCysLeuAlaSerCysAspSer 180

Qy    580 AAACGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTCTGTGCAATTGGA 639
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysArgThrAspValThrGlyAlaGlnIleAlaLeuAlaLeuAlaIleGlyPheSerValAlaIleGly 200
```

```
Qy   640  CATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCT  699
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaAspProLeuGlyPro  220

Qy   700  GCAGTTATCATGGGAAAATTGGGAAAACCATTGGATATATTGGGTGGGCCATCATAGGA  759
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221  AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly  240

Qy   760  GCTGTCCTCCTGGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGT  819
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  AlaValLeuAlaGlyGlyLeuTyrTyrTyrValPheCysProValGluAspValGlyLysArg  260

Qy   820  CGTTTAAGAAGCCTTCAGCAAACGCTGCCCAGCAAACAAAAGAAGCTACATGGAGGTG  879
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   261  ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy   880  GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCAT  939
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   281  GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis  300

Qy   940  GTGATTGACGTTGACCGGGAGAGAGAAGAGGGAAAGACCAATCTCGAGAGTATTG  999
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  ValIleAspValAspArgGluArgGluGluGlyLysThrAsnLeuGluSerGlyGluValLeu  320

Qy  1000  TCTTCAGTA  1008
          |||||||||
Db   321  SerSerVal  323
```

RESULT 2
AQP4_BOVIN
ID   AQP4_BOVIN   STANDARD;   PRT;   323 AA.
AC   O77750;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Aquaporin 4 (WCH4) (Mercurial-insensitive water channel) (MIWC).
GN   Name=AQP4; Synonyms=AQP-4;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=Holstein; TISSUE=Brain;
RX   MEDLINE=20135596; PubMed=10135596; DOI=10.1016/S0167-4781(99)00194-3;
RA   Sobue K., Yamamoto N., Yoneda K., Fujita K., Miura Y., Asai K.,
RA   Tsuda T., Katsuya H., Kato T.;
RT   "Molecular cloning of two bovine aquaporin-4 cDNA isoforms and their
RT   expression in brain endothelial cells.";
RL   Biochim. Biophys. Acta 1489:393-398(1999).
CC   -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC   regulates body water balance and mediates water flow within the
CC   central nervous system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=2; Synonyms=B;
CC   IsoId=O77750-1; Sequence=Displayed;
CC   Name=1; Synonyms=A;
CC   IsoId=O77750-2; Sequence=VSP_003231;
CC   -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC   three membrane-spanning domains and a pore-forming loop with the
CC   signature motif Asn-Pro-Ala (NPA).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --

CC   ---
DR   EMBL; AB015947; BAA36505.2; -.
DR   EMBL; AB012950; BAA33583.1; -.
DR   EMBL; AB028642; BAA89291.1; -.
DR   HSSP; P29972; 1FQY.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW   Transport.
FT   DOMAIN   1   36   Cytoplasmic (Potential).
FT   TRANSMEM   37   57   Potential.
FT   DOMAIN   58   64   Extracellular (Potential).
FT   TRANSMEM   65   85   Potential.
FT   DOMAIN   86   115   Cytoplasmic (Potential).
FT   TRANSMEM   116   136   Potential.
FT   DOMAIN   137   155   Extracellular (Potential).
FT   TRANSMEM   156   176   Potential.
FT   DOMAIN   177   184   Cytoplasmic (Potential).
FT   TRANSMEM   185   205   Potential.
FT   DOMAIN   206   231   Extracellular (Potential).
FT   TRANSMEM   232   252   Potential.
FT   DOMAIN   253   323   Cytoplasmic (Potential).
FT   SITE   97   99   NPA 1.
FT   SITE   213   215   NPA 2.
FT   MOD_RES   285   285   Phosphoserine (By similarity).
FT   CARBOHYD   153   153   N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   206   206   N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   1   22   Missing (in isoform 1).
FT   /FTId=VSP_003231.
SQ   SEQUENCE   323 AA; 34616 MW; E8E14507E6474D5F CRC64;

Alignment Scores:
Pred. No.:   Length: 323
Score: 1628.00   Matches: 311
Percent Similarity: 97.52%   Conservative: 4
Best Local Similarity: 96.28%   Mismatches: 8
Query Match: 74.64%   Indels: 0
DB: 1   Gaps: 0

US-10-723-180-1 (1-1152) x AQP4_BOVIN (1-323)
```

```
Qy    40  ATGAGTGACAGACCCACCAGCAAGGCGGTGGGGTAAGTGCGAAGTGGACCTTTGTGTACCAGAGAG   99
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MetSerAspAlaProAlaAlaThrArgTrpArgTrpGlyLysCysGlyProLeuCysThrArgGlu   20

Qy   100  AACATCATGTGGGCTTCAAAGGGGTCTGACTCAAGCTTTCTGGAAAGCAGTCACAGCG  159
          |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21  SerIleMetValAlaPheLysGlyValTrpValTrpProGlnThrPheTrpLysAlaValThrAla   40

Qy   160  GAATTCTGCCCATGCCATGCTTATTTTGTCTCCTCCAGCCTGGATCCACCATCAACTGGGGT  219
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41  GluPheLeuAlaMetLeuCysLeuLeuPheLeuValLeuLeuLeuSerLeuGlySerLeuAsnTrpGly   60

Qy   220  GGAACAGAAAAGCCCTTACCGGTCGACATGGTTCGACATGGTTGACTCAGC  279
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GlyAlaGluGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer   80

Qy   280  ATTGCAACCATGGTGCAGTGCTTGGCCATATCAGCGGTGGCCACATCACCCGTGCAGTG  339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal  100

Qy   340  ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCGAAGTCGTGCTTCTACATCGCA  399
          ||||||||||||||||||::::|||||||||||||||||||||||||||||||||||||||
Db   101  ThrValAlaMetValCysThrArgArgIleSerThrArgArgIleAlaSerValPheTyrIleAla  120

Qy   400  GCCCAGTGCCTGGGGCCATCATTGGAGCAGGAGAATCCTCATCTGTCACACCTCCCAGT  459
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer  140
```

```
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Porin; Transmembrane; Transport.
SQ   SEQUENCE   323 AA;  34754 MW;  D95929E21374948D CRC64;

Alignment Scores:
Pred. No.:                              Length:           323
Score:                  7.3e-115        Matches:          310
Percent Similarity:       97.52%        Conservative:       5
Best Local Similarity:    95.98%        Mismatches:         8
Query Match:              74.42%        Indels:             0
DB:                            2        Gaps:               0

US-10-723-180-1 (1-1152) x Q866S4 (1-323)

Qy   40 ATGAGTGACAGACCCACAGCAAGGCGGTGGGTAAGTGTGGACCTTTGTGTACCAGAGAG   99
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MetSerAspArgProAlaAlaArgArgTrpGlyLysCysCysThrArgGlu          20

Qy  100 AACATCATGGTGGCTTCAAAGGGGTCTGACTCCCAAGCTTTCTGGAAAGCAGTCACAGCG  159
           :::||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db   21 SerIleMetValAlaAlaPheLysGlyValTrpThrGlnThrPheThrLysAlaValThrAla  40

Qy  160 GAATTTCTGGCCATGCTTATTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGT  219
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerLeuLeuAsnTrpGly    60

Qy  220 GGAACAGAAAAGCCTTACCGGTCGACATGGTTCATCTCCCTTGCTTGGACTCAGC      279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GlyAlaGluLysProLeuProValAspMetValHisIleSerLeuGlyLeuSer        80

Qy  280 ATTGCAACCATGGTGCAGTGCTTGGCCATATCAGGCGGTGGCCAATCAACCCTGCAGTG  339
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal  100

Qy  340 ACTGTGGCCATGGTGTGACCAGGAGATCAGCATCATCGCCAAGTCTGTCTTCTACATCGCA  399
           ::::|||||||||||||||||||||||||||||||::::||||||||||||||||||||
Db  101 ThrValAlaMetValCysThrArgArgIleSerIleAlaLysSerIleAlaValPheTyrIleAla  120

Qy  400 GCCCAGTGCTCGGGGCCATCATTGGAGCAGGAGAATCCTATCTGGTCACACCTCCCAGT  459
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AlaGlnCysSerGlyLeuGlyAlaIleLeuGlyAlaGlyIleAlaLeuTyrLeuValThrProProSer  140

Qy  460 GTGGTGGGAGGCCTGGAGTCACCATGGTTCATGGAAAATTACCGCTGGTCATGGTCTC  519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 ValValGlyGlyLeuGlyValThrThrValHisArgAsnLeuSerAlaGlyHisGlyLeu  160

Qy  520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC  579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy  580 AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly  200

Qy  640 CATTATTGCAATCAATTATCTGGTGCCAGCATGAATCCGCCCGATCCTTTGGACCT  699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy  700 GCAGTTATCATGGGAAATTGGAAAACCATTGGATATATTGGGTTGGGCCCATCATAGGA  759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly  240

Qy  760 GCTGTCCCTGCCTGTGGGCCTTTATGAGTATGTCTTCGTCCAGATGTTGAATTCAAACGT  819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260
```

```
Qy  460 GTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGAAATCTTACCGCTGGTCATGGTCTC  519
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 ValValGlyGlyLeuGlyValThrThrValHisGlyAsnLeuSerAlaGlyHisGlyLeu  160

Qy  520 CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC  579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161 LeuValGluLeuIleIleLeuIlePheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy  580 AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTTCTGTTGCAATTGGA  639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly  200

Qy  640 CATTATTGCAATCAATTATCTGGTGCCAGCATGAATCCGCCCGATCCTTTGGACCT  699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy  700 GCAGTTATCATGGGAAATTGGAAAACCATTGGATATATTGGGTTGGGCCCATCATAGGA  759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly  240

Qy  760 GCTGTCCCTGCCTGTGGGCCTTTATGAGTATGTCTTCGTCCAGATGTTGAATTCAAACGT  819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy  820 CGTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGGAAGCTACATGGAGGTG  879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261 ArgPheLysLysGluAlaPheSerLysSerLeuProAlaAsnLysArgLysLeuHisGlyGlyVal  280

Qy  880 GAGGACAACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCAT  939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis  300

Qy  940 GTGATTGACGTTGACCGGGGAGGAGGAGAAGAAGGGGAAAAGACCAATCTGGAGAGGTATTG  999
           ||||||||||||||||||||||||||||||||     |||||||||||||||||||||||
Db  301 ValIleAspValAspArgGlyGlyGlyGluGluGlyGluLysThrAsnLeuGluArgTyr  320

Qy 1000 TCTTCAGTA 1008
           ||||||||
Db  321 SerSerVal 323
```

```
RESULT 3
Q866S4
ID   Q866S4      PRELIMINARY;    PRT;   323 AA.
AC   Q866S4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Aquaporin 4A (Aquaporin-4 M1 isoform).
GN   Name=AQP4A; Synonyms=AQP4;
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Coghlan M.W., Koukoulas I., Armgam A., Jeyaseelan K.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bondar A.A., Zelenina M.N., Gunnarson E., Illarionova N.B., Aperia A.,
RA   Zelenin S.M.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR   EMBL; AY177612; AAO21366.1; -.
DR   EMBL; AY304480; AAQ74771.1; -.
DR   HSSP; P29972; 1H6I.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019867; C:outer membrane; IEA.
DR   GO; GO:0015288; F:porin activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
```

```
Qy    820  CGTTTTAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGAGAGCTACATGAGGTG  879
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261  ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy    880  GAGGACAACAGGAGTCAGGTAGAGACGTAGAGACGGATGACCTGATTCTAAAACCTGAGTGTGCAT  939
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281  GluAspAsnArgSerGlnValGluThrArgAspAspLeuIleLeuLysProGlyValValHis  300

Qy    940  GTGATTGACGTTGACCCGGGAGAGGAAGAAGAGAAAGACCAATCTGGAGAGGTATTG  999
                ||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  ValIleAspIleAspArgArgGlyGluGluLysGlyLysAspProSerGlyGluValLeu  320

Qy   1000  TCTTCAGTA 1008
                ||||||
Db    321  SerSerVal 323

RESULT 4
AQP4_DIPME
ID   AQP4_DIPME     STANDARD;     PRT;     324 AA.
AC   Q92J4; Q92J5;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Aquaporin 4.
GN   Name=AQP4;
OS   Dipodomys merriami (Merriam's kangaroo rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Heteromyidae;
OC   Dipodomyinae; Dipodomys.
OX   NCBI_TaxID=94247;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fang P.K., Huang Y., Walsberg G.E., Brown D., van Hoek A.N.;
RT   "Molecular cloning and characterization of Merriam's kangaroo rat
RT   aquaporin 4.";
RL   Am. J. Physiol. 280:F794-F802(2001).
RN   [2]
RP   TISSUE SPECIFICITY.
RX   PubMed=11292621;
RA   Huang Y., Tracy R., Walsberg G.E., Makkinje A., Fang P., Brown D.,
RA   Van Hoek A.N.;
RT   "Absence of aquaporin-4 water channels from kidneys of the desert
RT   rodent Dipodomys merriami merriami.";
RL   Am. J. Physiol. 280:F794-F802(2001).
CC   -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC       regulates body water balance and mediates water flow within the
CC       central nervous system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=2;
CC         IsoId=Q92J4-1; Sequence=Displayed;
CC       Name=1;
CC         IsoId=Q92J4-2; Sequence=VSP_010210;
CC   -!- TISSUE SPECIFICITY: Not expressed in kidney, Detectable in gastric
CC       parietal and brain astroglial cells. The absence of AQP4 in kidney
CC       may be critical for the extreme urinary concentration that occurs
CC       in this species (up to 5,000 mosmol/kg H(2)O).
CC   -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC       three membrane-spanning domains and a pore-forming loop with the
CC       signature motif Asn-Pro-Ala (NPA).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AY032857; AAK66823.1; -.
DR   EMBL; AY032858; AAK66824.1; -.
DR   HSSP; P29972; 1FQY.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW   Transport.
FT   DOMAIN        1     36       Cytoplasmic (Potential).
FT   TRANSMEM     37     57       Potential.
FT   DOMAIN       58     63       Extracellular (Potential).
FT   TRANSMEM     64     84       Potential.
FT   DOMAIN       85    115       Cytoplasmic (Potential).
FT   TRANSMEM    116    136       Potential.
FT   DOMAIN      137    155       Extracellular (Potential).
FT   TRANSMEM    156    176       Potential.
FT   DOMAIN      177    184       Cytoplasmic (Potential).
FT   TRANSMEM    185    205       Potential.
FT   DOMAIN      206    231       Extracellular (Potential).
FT   TRANSMEM    232    252       Potential.
FT   DOMAIN      253    324       Cytoplasmic (Potential).
FT   SITE         97     99       NPA 1.
FT   SITE        213    215       NPA 2.
FT   MOD_RES     285    285       Phosphoserine (By similarity).
FT   CARBOHYD    153    153       N-linked (GlcNAc...) (Potential).
FT   CARBOHYD    206    206       N-linked (GlcNAc...) (Potential).
FT   VARSPLIC      1     22       Missing (in isoform 1).
FT                               /FTId=VSP_010210.
SQ   SEQUENCE   324 AA;  34939 MW;  EAE9CE3B17952CDD CRC64;

Alignment Scores:
Pred. No.:
Score:                    8.25e-113          Length:          324
                          1596.00            Matches:         304
Percent Similarity:       96.87%             Conservative:      5
Best Local Similarity:    95.30%             Mismatches:       10
Query Match:              73.18%             Indels:            0
DB:                       1                  Gaps:              0

US-10-723-180-1 (1-1152) x AQP4_DIPME  (1-324)

Qy     40  ATGAGTGACAGACCCCACAGCAGGCCGTGGGTAAGTGCGGTAGTGCGACCTTTGTGTACCAGAGAG  99
                |||||||||||||||||||||| ||| |||||||||||||||||||
Db      1  MetSerAspArgProAlaAlaArgProTrpGlyLysCysGlySerLeuCysArgArgGlu  20

Qy    100  AACATCATGGTGGCTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCCG  159
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21  GluIleMetValAlaPheLysGlyValThrGlyAlaPheThrLysSerAlaValThrAla  40

Qy    160  GAATTTCTGGCCATGGTCTTATTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGT  219
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41  GluPheLeuAlaMetLeuPheValLeuLeuPheValLeuLeuSerLeuGlySerThrIleAsnTrpGly  60

Qy    220  GGAACAGAAAAGCCTTACCGGTGCGACATGGTTCTCATCTCCCTTGCTTGCTTGGACTCAGC  279
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GlyLysGluAsnProLeuProValAspMetValLeuIleSerLeuAlaCysPheGlyLeuSer  80

Qy    280  ATTGCAACCATGGTGCAGTGCTTTGGCCATATCAGCGGTGGCCACATATCGCATCCTGCAGTG  339
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleSerHisProAlaVal  100

Qy    340  ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATGGCCAAGTCTGTTCTTCTACATCGCA  399
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101  ThrValAlaMetValCysThrArgLysIleSerMetAlaLysSerValPheTyrIleAla  120

Qy    400  GCCCAGTGCCTGGGGCCCATCATTGGAGCAGCAGGAATCCTCATCCTCATCTGGTCACACCTCCCAGT  459
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AlaGlnCysLeuGlyLysAlaIleLeuGlyAlaAlaGlyIleLeuIleLeuIleTrpLeuValThrProProSer  140

Qy    460  GTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCGGTGGTCATGGTCTC  519
```

```
Db        |||||||||||||||||||||||  |||||||||||  ||||||||||||||||||  ||||
     141  ValValGlyGlyLeuGlyValThrThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu  160

Qy   520  CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC       579
          |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db   161  LeuValGluLeuIleIleThrPheGlnLeuLeuValPheThrIlePheAlaSerCysAspSer   180

Qy   580  AAACGGACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGA     639
          |||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db   181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly     200

Qy   640  CATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCGCCTTTGGACCT       699
          ||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db   201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaAlaPheGlyProProPro  220

Qy   700  GCAGTTATCATGGGAAATGGGAAAACCATTGGATATATTGGGGTGGGCCATCATAGA       759
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   221  AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly     240

Qy   760  GCTGTCCTCGCTGGTGGCCTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGT     819
          ||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db   241  AlaValLeuAlaGlyGlyLeuMetSerMetSerSerValGlnMetLeuAsnSerAsnVal    260

Qy   820  CGTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATGGAGGTG     879
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   261  ArgPheLysGluAlaPheSerLysLeuAlaAlaGlnGlnThrLysGlySerTyrMetGluVal  280

Qy   880  GAGGACAACAGGAGTCAGGTAGAGACGACGGATGACCTGACTTCTAAAACCTGGAGTGGTGCAT  939
          |||||||||||||||||||||||||||||||||  ||||||||||||
Db   281  GluAspAsnArgSerGlnValGluThrGlnAspAspLeuGlnLysProGlyLeuValHis     300

Qy   940  GTGATTGACGTTGACCCGGGAGGAGGAAGAAGAGGGGAAAGACCAATCTCGAGAGGTA     996
          |||||||||:::||||||||:::|||||||||||:::|||||||||||||||:::
Db   301  ValIleAspIleAspProGlyArgArgAspGlyGluArgGlyLysLysAspProSerGlyGluIle  319
```

```
RESULT 5
AQP4_RAT
ID   AQP4_RAT      STANDARD;      PRT;     323 AA.
AC   P47863;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Aquaporin 4 (WCH4) (Mercurial-insensitive water channel) (MIWC).
GN   Name=Aqp4;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF HIS-201.
RC   TISSUE=Brain;
RX   MEDLINE=95108097; PubMed=7528931;
RA   Jung J.S., Bhat R.V., Preston G.M., Guggino W.B., Baraban J.M.,
RA   Agre P.;
RT   "Molecular characterization of an aquaporin cDNA from brain: candidate
RT   osmoreceptor and regulator of water balance.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:13052-13056(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Chen D., Chen J., Jing K., Simon R.P., Graham S.H.;
RT   "Isolation of an aquaporin-4 water channel (AQP4) gene induced
RT   following cerebral ischemia from the rat brain using modified
RT   subtractive hybridization and differential screening.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 23-323 FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=94164885; PubMed=7509789;
RA   Hasegawa H., Ma T., Skach W., Matthay M.A., Verkman A.S.;
RT   "Molecular cloning of a mercurial-insensitive water channel expressed
RT   in selected water-transporting tissues.";
RL   J. Biol. Chem. 269:5497-5500(1994).
```

```
RN   [4]
RP   PHOSPHORYLATION SITE SER-285.
RX   PubMed=12692561; DOI=10.1038/nbt819;
RA   Wu C.C., MacCoss M.J., Howell K.E., Yates J.R. III;
RT   "A method for the comprehensive proteomic analysis of membrane
RT   proteins.";
RL   Nat. Biotechnol. 21:532-538(2003).
CC   -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC       regulates body water balance and mediates water flow within the
CC       central nervous system. It is expressed predominantly in the
CC       ependymal cell lining the aqueductal system and over the space of
CC       the brain in contact with the subarachnoid space, as cerebrospinal
CC       fluid fills these structures it may facilitate water balance
CC       between brain parenchyma and the fluid compartment. In the plasma
CC       membranes of the neurons of the paraventricular and supraoptic
CC       nuclei, it may mediate rapid changes in cell volume in response to
CC       local shifts in extracellular osmolality.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P47863-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P47863-2; Sequence=VSP_003235;
CC   -!- TISSUE SPECIFICITY: Abundant in mature brain but only weakly
CC       detectable in eye, kidney, intestine, and lung.
CC   -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC       three membrane-spanning domains and a pore-forming loop with the
CC       signature motif Asn-Pro-Ala (NPA).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-23 is the initiator.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; U14007; AAC52152.1; -.
DR   EMBL; AF144082; AAD37965.1; -.
DR   EMBL; L27588; AAA17730.1; -.
DR   PIR; I59283; I59283.
DR   HSSP; P29972; 1FQY.
DR   RGD; 2143; Aqp4.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW   Transport.
FT   DOMAIN        1     36       Cytoplasmic (Potential).
FT   TRANSMEM     37     57       Potential.
FT   DOMAIN       58     64       Extracellular (Potential).
FT   TRANSMEM     65     85       Potential.
FT   DOMAIN       86    115       Cytoplasmic (Potential).
FT   TRANSMEM    116    136       Potential.
FT   DOMAIN      137    155       Extracellular (Potential).
FT   TRANSMEM    156    176       Potential.
FT   DOMAIN      177    184       Cytoplasmic (Potential).
FT   TRANSMEM    185    205       Potential.
FT   DOMAIN      206    231       Extracellular (Potential).
FT   TRANSMEM    232    252       Potential.
FT   DOMAIN      253    323       Cytoplasmic (Potential).
FT   SITE         97     99       NPA 1.
FT   SITE        213    215       NPA 2.
FT   MOD_RES     285    285       Phosphoserine.
FT   CARBOHYD    153    153       N-linked (GlcNAc... ) (Potential).
FT   VARSPLIC    150    204       Missing (in isoform Short).
FT                               /FTId=VSP_003235.
```

```
FT   MUTAGEN    201   201       H->P:  Partial loss of transport activity.
FT   CONFLICT   201   201       H -> P (in Ref. 3).
SQ   SEQUENCE   323 AA;   34480 MW;   6ADD246477I3609D CRC64;

Alignment Scores:
Pred. No.:                                   323
Score:                  9.83e-113            304
Percent Similarity:     1595.00              8
Best Local Similarity:  96.59%               11
Query Match:            94.12%               0
DB:                     73.13%               0

US-10-723-180-1 (1-1152) x AQP4_RAT (1-323)
```

```
Qy    40  ATGAGTGACAGACCCACAGCAAGGCGGTGGGTAAGTGTGGGACCTTTGTGTACCAGAGAG   99
              ||||||||||||||||||||||||||||||||||||||||||   |||:::|||||||||
Db     1  MetSerAspGlyAlaAlaAlaArgArgTrpGlyLysCysGlyProCysSerArgGlu    20

Qy   100  AACATCATGGTGGCCTTTCAAAGGGGTCTGGACTTCAAGCTTCTTGGAAAGCAGTCACAGCG  159
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21  SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla   40

Qy   160  GAATTTCTGGCCATGCTTATTTTTGTCTCCTCAGCCTGGATCAACTCACCATCAACTGGGGT  219
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41  GluPheLeuAlaMetLeuIlePheValLeuSerValGlySerThrIleAsnTrpGly   60

Qy   220  GGAACAGAAAAGCCTTACCGGTGCACCAGTGGTTCTCATCTCCCTTGCTTGGACTCAGC  279
              ||||:::  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer   80

Qy   280  ATTGCAACCATGGTGCAGTGCTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG  339
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyHisIleSerGlyProAlaVal  100

Qy   340  ACTGTGGCCATGGTGTGCACCAGAAGATCAGCATCGCCAAGTCGTCTTCTACATCGCA  399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101  ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleThr  120

Qy   400  GCCCAGTGCCTGGGGGCCATCATTGGAGCAGGAATCCTCATCGGTCACACCTCCCAGT  459
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuIleGlyValThrProProSer  140

Qy   460  GTGGTGGGAGGCCTGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC  519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141  ValValGlyGlyLeuGlyValThrThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu  160

Qy   520  CTGGTTGAGTTGATAATCACATTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC  579
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161  LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy   580  AAACGGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTCTTGTTGCAATTGGA  639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  LysArgThrAspValThrGlySerValAlaLeuAlaIleGlyPheSerValAlaIleGly  200

Qy   640  CATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCACCTTTGGACCT  699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy   700  GCAGTTATCATGGGAAAATGGAAAACCATTGGATATATTGGGGTTGGGCCCATCATAGGA  759
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221  AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly  240

Qy   760  GCTGTCCTCGTGGCCTTTATGTATGTCTTCTGTCCAGATGTTGAATTCAAACGT  819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  AlaValLeuAlaGlyAlaLeuTyrGlyTyrValPheCysProAspValGluLeuLysArg  260

Qy   820  CGTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAAGGAAGCTACATCGGAGGTG  879
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   261  ArgLeuLysLysGluLeuProThrSerLysGlnGlnGluAlaArgLeuGlySerTyrMetGluVal  280

Qy   880  GAGGACAACAGGAGTCAGTAGACACGATGACCTGATTCTAAAACCTGGAGTTGGTGCAT  939
              |||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db   281  GluAspAsnArgSerGlnValGlnThrGluValAspAspMetProLeuIleLeuLysProGlyValValHis  300
```

```
Qy   940  GTGATTGACGTTGACCGGGGAGAGAGAGAAGGGGAAAGACCAATCTGGAGAGGTATTG  999
              |||||||||||||||||||||:::|||||||||||:::||||||||||||||||||||
Db   301  ValIleAspValAspArgGlyAspGlyGluArgGluGlyGluLysGlyGlyLysGlyProSerSerGlyGluValLeu  320

Qy  1000  TCTTCAGTA  1008
              ||||||||
Db   321  SerSerVal  323
```

```
RESULT 6
AQP4_MOUSE
ID   AQP4_MOUSE    STANDARD;   PRT;   323 AA.
AC   P55088; P97818; Q61131; Q61132; Q8VHE4; Q8VHE5; Q9EQI3;
DT   01-OCT-1996 (Rel. 34, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Aquaporin 4 (WCH4) (Mercurial-insensitive water channel) (MIWC).
GN   Name=Aqp4;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=96299634; PubMed=8660998; DOI=10.1006/geno.1996.0214;
RA   Ma T., Yang B., Verkman A.S.;
RT   "Gene structure, cDNA cloning, and expression of a mouse mercurial-
RT   insensitive water channel.";
RL   Genomics 33:382-388(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2)
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=97288526; PubMed=9143504; DOI=10.1006/geno.1997.4641;
RA   Turtzo L.C., Lee M.D., Lu M., Smith B.L., Copeland N.G., Gilbert D.J.,
RA   Jenkins N.A., Agre P.;
RT   "Cloning and chromosomal localization of mouse aquaporin 4: exclusion
RT   of a candidate mutant phenotype, ataxia.";
RL   Genomics 41:267-270(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT ARG-4.
RX   MEDLINE=20419017; PubMed=10960499;
RA   Zelenin S., Gunnarson E., Alikina T., Bondar A., Aperia A.;
RT   "Identification of a new form of AQP4 mRNA that is developmentally
RT   expressed in mouse brain.";
RL   Pediatr. Res. 48:335-339(2000).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC       regulates body water balance and mediates water flow within the
CC       central nervous system.
```

```
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=2; Synonyms=MIWC2, AQP4-M1;
CC         IsoId=P55088-1; Sequence=Displayed;
CC       Name=1; Synonyms=MIWC1, AQP4-M21;
CC         IsoId=P55088-2; Sequence=VSP_003233;
CC       Name=3; Synonyms=MIWC3;
CC         IsoId=P55088-3; Sequence=VSP_003234;
CC   -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC       three membrane-spanning domains and a pore-forming loop with the
CC       signature motif Asn-Pro-Ala (NPA).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U48398; AAB41569.1; -.
DR   EMBL; U48397; AAB41568.1; -.
DR   EMBL; U33012; AAA84923.1; -.
DR   EMBL; U48400; AAB41571.1; -.
DR   EMBL; U48399; AAB41570.1; -.
DR   EMBL; U88623; AAC53155.1; -.
DR   EMBL; AF469168; AAL73545.1; -.
DR   EMBL; AF469169; AAL73546.1; -.
DR   EMBL; AF219992; AAG44243.2; -.
DR   EMBL; BC024526; AAH24526.1; -.
DR   HSSP; P29972; 1FQY.
DR   MGD; MGI:107387; Aqp4.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   GO; GO:0015250; F:water channel activity; IDA.
DR   GO; GO:0030104; P:water homeostasis; NAS.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW   Transport.
FT   DOMAIN        1     36       Cytoplasmic (Potential).
FT   TRANSMEM     37     58       Potential.
FT   DOMAIN       59     64       Extracellular (Potential).
FT   TRANSMEM     65     85       Potential.
FT   DOMAIN       86    115       Cytoplasmic (Potential).
FT   TRANSMEM    116    136       Potential.
FT   DOMAIN      137    155       Extracellular (Potential).
FT   TRANSMEM    156    176       Potential.
FT   DOMAIN      177    184       Cytoplasmic (Potential).
FT   TRANSMEM    185    205       Potential.
FT   DOMAIN      206    231       Extracellular (Potential).
FT   TRANSMEM    232    252       Potential.
FT   DOMAIN      253    323       Cytoplasmic (Potential).
FT   SITE         97     99       NPA 1.
FT   SITE        213    215       NPA 2.
FT   MOD_RES     285    285       Phosphoserine (By similarity).
FT   CARBOHYD    153    153       N-linked (GlcNAc...) (Potential).
FT   CARBOHYD    206    206       N-linked (GlcNAc...) (Potential).
FT   VARSPLIC      1     22       Missing (in isoform 1).
FT                                /FTId=VSP_003233.
FT   VARSPLIC      1      1       MSDGAAARRWG -> MVHGFGCFVFFFLISLSLWASEDST
FT                                CNSTLPLCHLATTLDCC (in isoform 3).
FT                                /FTId=VSP_003234.
FT   VARIANT       4      4       G -> R.
FT   CONFLICT     37     37       A -> S (in Ref. 1; AAB41571).
FT   CONFLICT     51     51       Missing (in Ref. 1).
FT   CONFLICT     88     88       F -> L (in Ref. 1).
FT   CONFLICT    174    174       I -> V (in Ref. 1).
```

```
FT   CONFLICT    313    313       K -> R (in Ref. 1).
SQ   SEQUENCE    323 AA;  34436 MW;  9AC638A3C3F412E2 CRC64;

Alignment Scores:
Pred. No.:                              Length:              323
Score:                  1.32e-110       Matches:             300
Percent Similarity:     1567.00         Conservative:          8
Best Local Similarity:  95.36%          Mismatches:            8
Query Match:            92.88%          Indels:                0
DB:                     71.85%          Gaps:                  0

US-10-723-180-1 (1-1152) x AQP4_MOUSE (1-323)

Qy     40 ATGAGTGACAGACCCACAGCAAGGCGGTGGGTAAGTGTGGACCTTTGTGTACCAGAGAG       99
          |||||||||                              ||:||||||||||
Db      1 MetSerAspGlyAlaAlaAlaArgArgTrpGlyHisSerGlyHisSerCysSerArgGlu   20

Qy    100 AACATCATGTGGCTTCAAAGGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCG     159
          :::||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheThrLysAlaValSerAla   40

Qy    160 GAATTTCTGGCCATGCTTATTTTTGTTCCCTCAGCCTGGGATCCAACCATCAACTGGGGT    219
          ||||||                              |||||||||
Db     41 GluPheLeuAlaThrLeuIlePheValLeuLeuValGlyValGlySerThrIleAsnTrpGly   60

Qy    220 GGAACAGAAAAGCCTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGC     279
          ||:::|||||||                              |||||
Db     61 GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer   80

Qy    280 ATTGCAACCATGGTGCAGTGCTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG    339
          ||||||||||                              |||||
Db     81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyValHisIleAsnProAlaVal  100

Qy    340 ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCA    399
          ||||||                              |||||
Db    101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleIle  120

Qy    400 GCCCAGTGCCTGCGGGGCCATCATTGGAGCAGGAGAATCCTATCTGGTCACACTTCCCAGT    459
          |||||||                              |||||
Db    121 AlaGlnCysLeuArgGlyHisHisTrpSerArgArgIleLeuTyrLeuValThrProProSer  140

Qy    460 GTGGTGGAGGCCTGGAGTCACCATGGTTCATGGAAAATCTTACCGCTGGTCATGGTCTC    519
          ||||||                              |||||
Db    141 ValValGlyGlyLeuGlyValThrThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu  160

Qy    520 CTGGTTGAGTGATAATCACATTTCAATTGGTGTTACTATCTTTGCCAGCTGTGATTCC    579
          ||||||                              |||||
Db    161 LeuValGluLeuIleLeuThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer  180

Qy    580 AAACGGACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCAATTGGA    639
          ||||||                              |||||
Db    181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly  200

Qy    640 CATTATTTGCAATCAATATACTGGTGCCAGCATGAATCCGCCCGATCCTTTGGACCT    699
          ||||||                              |||||
Db    201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro  220

Qy    700 GCAGTTATCATGATGGGAAAATTGGGAAAACCATTGGATATATTGGGTGGCCCATCATAGGA    759
          ||||||                              |::::|||
Db    221 AlaValIleMetMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGly  240

Qy    760 GCTGTCCTGCTGCTGGCCTTTATGAGTATGTCTTCGTCCAGATGTTGAATTCAAACGT    819
          ||||||                              |||||
Db    241 AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArg  260

Qy    820 CGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGAAGAAGTACATGGAGGTG    879
          |||                              ||||||
Db    261 ArgLeuLysGluAlaPheSerLysAlaLeuGlnGlnThrLysLysLysTyrMetGluVal  280

Qy    880 GAGGACAACAGGAGTCAGGTAGAGACGGATGAACCTGATTCTAAAACCTGCAGTGTGCAT    939
          ||||||                              |::::||
Db    281 GluAspAsnArgSerGlnValGluThrAspGluProAspSerLysThrCysSerValHis  300
```

Db   81 AlaMetValCysThrArgArgIleSerIleAlaAlaValAlaValPheTyrIleAlaAlaAlaGln 100

Qy 406 TGCTCGGGGGCCATCATTGGACCAGGAATCCTATCTGGTCACACCTCCCAGTGTCGTG 465
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 CysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProSerValVal 120

Qy 466 GGAGGCGCTGGAGTCACCATGGTTCATGGAAATCTTACCGCGTCGTCATGGTCTCCTGGTT 525
      ||||||||||||||||||||||||||||||| :::||||||||||||||||||
Db 121 GlyGlyLeuGlyValThrThrValHisGlyAsnLeuSerAlaGlyHisGlyLeuLeuVal 140

Qy 526 GAGTTGATAATCACATTCAATTGGTGTTTACTACTCTTTGGCCAGCTGTGATTCCAAACGG 585
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 GluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArg 160

Qy 586 ACTGATGTCACTGGCTCCAATAGCTTTAGCAATTGGATTTCTGTTGCAATTGGACATTA 645
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 ThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGlyHisLeu 180

Qy 646 TTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCCGATCCTTTGGACCTGCAGTT 705
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaVal 200

Qy 706 ATCATGGGAAATTGGAAAACCATTGGATATATTGGTTGGGCCCATCATAGGAGCTGTC 765
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 IleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaVal 220

Qy 766 CTCGCTGGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTT 825
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 LeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArgArgPhe 240

Qy 826 AAAGAGCCTTCAGCAAAGTCGCCCAGCAAACAAAAGGAAGCTACATGGAGGTGGAGGAC 885
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 LysGluProSerAlaLysSerAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAsp 260

Qy 886 AACAGGAGTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTGGTGCATGTGATT 945
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 261 AsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHisValIle 280

Qy 946 GACGTTGACCGGGGAGAGGAGAAGAGGGAAAGACCAATCTGGAGAGGTATTGTCTTCA 1005
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 281 AspIleAspArgGlyGluGluLysGlyLysGlyLysGlyGluAspProSerGlyValLeuSerSer 300

Qy 1006 GTA 1008
      |||
Db 301 Val 301

RESULT 8
Q8UVB8
ID Q8UVB8   PRELIMINARY;   PRT;   335 AA.
AC Q8UVB8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Aquaporin-4.
GN Name=AQP4;
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Cui Y., Yang Y., Bolan J., Wang W., Nishimura H., Fan Z.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AF465730; AAL73511.1; -.
DR HSSP; P29972; 1IH5.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.

Qy 940 GTGATTGACGTTGACCGGGGAGAGGAGAAGAGGGAAAGACCAATCTGGAGAGTATTG 999
      ||||||::|||||||||||||||||||||||||||||||||||||| |||||||||
Db 301 VallleAspIleAspArgGlyGluGluLysGlyLysGlyLysGlyGluAspProSerGlySerGlyGluValLeu 320

Qy 1000 TCTTCAGTA 1008
      |||||||||
Db 321 SerSerVal 323

RESULT 7
Q6XVT6
ID Q6XVT6   PRELIMINARY;   PRT;   301 AA.
AC Q6XVT6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Aquaporin 4 M23 isoform.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Bondar A.A., Zelenin S., Gunnarson E., Aperia A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AY206465; AAO38843.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00783; MINTRINSICP.
DR ProDom; PD000295; MIP; 1.
DR TIGRFAMs; TIGR00861; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Porin; Transmembrane; Transport.
SQ SEQUENCE 301 AA; 32282 MW; 36417087A5B2AFE9 CRC64;

Alignment Scores:
Pred. No.:   6.82e-106   Length: 301
Score: 1505.00   Matches: 290
Percent Similarity: 97.67%   Conservative: 4
Best Local Similarity: 96.35%   Mismatches: 7
Query Match: 69.01%   Indels: 0
DB: 2   Gaps: 0

US-10-723-180-1 (1-1152) x Q6XVT6 (1-301)

Qy 106 ATGGTGGCTCTTCAAAGGGGTCCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCCGGAATTT 165
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetValAlaPheLysGlyValTrpThrGlnThrPheGlnThrPheThrArgLysAlaValThrAlaGluPhe 20

Qy 166 CTGGCCATGCTTATTTTTGTTCTCCCTCAGCCTGGGATCCACCATCAACTGGGGTGGAACA 225
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 21 LeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlyAla 40

Qy 226 GAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTGCA 285
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41 GluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla 60

Qy 286 ACCATGGTGCCAGTGCTTTGGCCACCATATCAGCGGTGGCCACCATACAACCCTGCAGTGACTGTG 345
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 ThrMetValGlnCysPheGlyHisGlyHisGlySerGlyGlyHisIleAsnProAlaValThrVal 80

Qy 346 GCCATGGTGCTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATCGCACGCCCAG 405
      |||||||||||||||||||||||||||||||||:::|||||||||||||||||

```
DR   GO; GO:0006810; P:transport; IEA.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Porin; Transmembrane; Transport.
SQ   SEQUENCE   335 AA;  36164 MW;  F942D4DA6F4DE8DA CRC64;

Alignment Scores:
Pred. No.:                         Length:          335
Score:               6.37e-103     Matches:         276
Percent Similarity:  92.79%        Conservative:    20
Best Local Similarity: 86.52%      Mismatches:      23
Query Match:         67.22%        Indels:          0
DB:                                Gaps:            0

US-10-723-180-1 (1-1152) x Q8UVB8 (1-335)

Qy     52 CCCACAGCAAGGCCGTGGGTAAGTGTGGACCTTGTGTACCAGAGAACATCATGGTG 111
             |||||| |||||       ||||||  |||||||||||||||||||  || |||
Db     17 ProProAlaArgSerSerLysGlyArgLeuCysLysGluAsnIleMetVal 36

Qy    112 GCTTTCAAAGGGGTCGGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCC 171
             |||||||| ||    |||||    |||||||:::||||||||||||||||||||||
Db     37 AlaPheLysGlyValTrpPheThrGlnProPheProPlysAlaValSerAlaGluPheLeuAla 56

Qy    172 ATGCTTATTTTGTTCTCCTCAGCCTGGGATCCACCATCAACTGGGGTCGAACAGAAAAG 231
             ||||||||||||||| ||||||||:::||||||||||||:::|||||:|||||
Db     57 MetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlySerGluLys 76

Qy    232 CCTTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATG 291
             |||||||||||||   ||||||  ||||||  ||||||  |||||||||||||||||
Db     77 ProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMet 96

Qy    292 GTGCAGTGCTTTGGCCATATCAGCGGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATG 351
             |||||||||| |||||||||||||||||| ||| ||:::|||||||  ||||||||||
Db     97 ValGlnCysPheGlyHisIleSerGlyValProSerValValProAlaValThrValAlaMet 116

Qy    352 GTGTGCACCAGGAGAATCAGCATCGCCAAGTGTTCTTCTACATCGACGCCCAGTGCCTG 411
             ||||||||||||||||||||  ||||| |||:::|||||||||||||||||||||
Db    117 ValCysThrArgLysIleSerLeuAlaLysSerLeuAlaLysPheTyrIleLeuAlaGlnCysLeu 136

Qy    412 GGGGCCATCATTGGAGCAGGAATCCTCTATCTCGTCACACCTCCCAGTGTGGTGGGAGGC 471
             ||||||||||| ||||||||||||| ||  |||||||||||||||  |||||| ||||
Db    137 GlyAlaIleValGlyAlaGlyIleLeuTyrLeuValThrProSerValValGlyGly 156

Qy    472 CTGGGAGTCACCATGGTTCATGGAAATCTTACCCGCTGGTCATGGTCTCCTGGTTGAGTTG 531
             ||||||||||||||||||||||:::|| |||||||||| |||||||||||||||||
Db    157 LeuGlyValThrAlaValHisGlyAsnLeuThrArgTrpSerTrpSerLeuLeuGluValGluLeu 176

Qy    532 ATAATCACATTTCAATTGGTTGTTTACTATCTTTGCCAGCTGTGATTCCAAACCGACTGAT 591
             || || ||||||| ||| |||:::|||||||||  |||||||||||||:||
Db    177 IleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysProThrAsp 196

Qy    592 GTCACTGGCTCAATAGCTTTAGCAATTGGATTTCGTGTGCAATTGGACATTTATTGCA 651
             |||||  |||| |  |||  ||||||||||||||  |||  ||||| |||||||||
Db    197 ValThrGlySerValAlaLeuAlaLeuAlaIleGlyPheArgValGlnLeuAspIlePheAla 216

Qy    652 ATCAATTATACTGTGCCAGTTTGGCCATATCAATCCCGGATCCTTTGACCTGCAGTTATCATG 711
             |||||||| |||||||  ||||||:::|||||||| |||||||||||||||||||||
Db    217 IleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleIleMet 236

Qy    712 GGAAATTGGGAAAACCATTGGATATATTGGGGTTGGGCCCATCATAGGAGAGTCGTTCCTGCT 771
             ||||||| ||||  |:::|| ||||||||  |||||| ||  ||||||||||||||
Db    237 GlyAsnTrpGluAsnGlnTrpValTyrTrpValTyrTrpValGlyProIleIleGlyGluAlaValAlaLeuAla 256

Qy    772 GGTGGCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAA 831
             ||||||:::  ||||||||||||||||||||:::||||||||||:||:||||
Db    257 GlyAlaLeuTyrGluTyrValTyrCysProAspValGluLeuLysArgArgPheLysAspAsp 276

Qy    832 GCCTTCAGCAACGCTGCCCAGCAACAAAAGGAAGCTACAGGGAGGTGAGGACAACAGG 891
```

next column

```
Db    277 ValPheSerLysThrSerGlnProSerLysGlyLysTyrIleGluValAspAspThrArg 296

Qy    892 AGTCAGGTAGAGACGGACGGACTCGATTCTAAAACCTGGAGTGGTGCATGTGATGATGACGTT 951
             |||||||||||||||||||||||||||||:::||||||||||||||||:::::||
Db    297 SerHisValGluThrAspAspAspLeuIleLeuLysProGlyIleValIleHisValIleAspIle 316

Qy    952 GACCCGGGAGGAGAGGAAGAGGGGAAAGACCAATCTGGAGAGTATTGTTCTTCAGTA 1008
             ||| |||:::||||||||||:::|| |||||
Db    317 AspArgSerGluAspLysGlyLysGlyArgAspProSerGluValLeuSerSerVal 335

RESULT 9
Q65YQ3
ID   Q65YQ3      PRELIMINARY;        PRT;    335 AA.
AC   Q65YQ3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Aquaporin 4.
GN   Name=AQP4;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Saito N., Ikegami H., Shimada K.;
RT   "cDNA cloning and mRNA expression of aquaporin 4 (AQP 4) in the
RT   chicken (Gallus domesticus).";
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB190358; BAD46731.1; -.
KW   Porin.
SQ   SEQUENCE   335 AA;  36258 MW;  66AE0B47CE75F6E5 CRC64;

Alignment Scores:
Pred. No.:                         Length:          335
Score:               1.25e-101     Matches:         273
Percent Similarity:  91.85%        Conservative:    20
Best Local Similarity: 85.58%      Mismatches:      26
Query Match:         66.44%        Indels:          2
DB:                                Gaps:            0

US-10-723-180-1 (1-1152) x Q65YQ3 (1-335)

Qy     52 CCCACAGCAAGGCCGTGGGTAAGTGTGGACCTTGTGTACCAGAGAACATCATGGTG 111
             |||||| |||||       ||||||  ||||||||||||||||:::|| |||
Db     17 ProProAlaArgSerSerLysGlyArgLeuCysLysGluSerIleMetVal 36

Qy    112 GCTTTCAAAGGGGTCGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCC 171
             |||||||| |||   |||||    |||||||:::||||||||||||||||||||||
Db     17 ProProAlaArgSerSerLysGlyAlaArgLeuCysLysGluSerIleMetVal 36

Qy    112 GCTTTCAAAGGGGTCTGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCC 171
             |||||||| |||   |||||    |||||||:::||||||||||||||||||||||
Db     37 AlaPheLysGlyValTrpProPheThrHisProPheProTrpPlysAlaValSerAlaGluPheLeuVal 56

Qy    172 ATGCTTATTTTGTTCTCCTCCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAG 231
             ||||||||||||||| ||||||||:::||||||||||||:::|||||:|||||
Db     57 MetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlySerGluLys 76

Qy    232 CCTTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATG 291
             |||||||||||||   ||||||  ||||||  ||||||  ||||||||||||||||
Db     77 ProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMet 96

Qy    292 GTGCAGTGCTTTGGCCATATCAGCGGGTGGGTAAGTGTGTACCAGAGAACATCATGGTG 111
             |||||||||||||||||||||||:::||||| |||||||||||||||:::| ||||
Db     17 ProProAlaArgSerSerLysGlyArgLeuCysLysGluSerIleMetVal 36

Qy    112 GCTTTCAAAGGGGTCTGACTCAAGCTTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCC 171
             |||||||| ||||||| ||||:::|| |||||||||||||||||||||||:::|||
Db     37 AlaPheLysGlyValTrpProPheHisProTrpProPheTrpProLysAlaValSerAlaGluPheLeuVal 56

Qy    172 ATGCTTATTTTGTTCTCCTCCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAG 231
             ||||||||||||||| ||||||||:::||||||||||||:::|||||:|||||
Db     57 MetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlySerGluLys 76

Qy    232 CCTTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATG 291
             |||||||||||||   ||||||  ||||||  ||||||  ||||||||||||||||
Db     77 ProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMet 96

Qy    292 GTGCAGTGCTTTGGCCATATCAGCGGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATG 351
             |||||||||| |||||||||||||||||| ||| ||:::|||||||  ||||||||||
Db     97 ValGlnCysPheGlyHisIleSerGlyValProSerValValProAlaValThrValAlaMet 116

Qy    352 GTGTGCACCAGGAGAGATCAGCATCGCCAAGTGTTCTTCTACATCGACGCCCAGTGCCTG 411
             |||||||||||||| |||||||  ||||| |||:::|||||||||||||||||||
Db    117 ValCysThrArgLysIleSerLeuAlaLysSerLeuAlaLysPheTyrIleLeuAlaGlnCysLeu 136

Qy    412 GGGGCCATCATTGGAGCAGGAATCCTCTATCTCGTCACACCTCCCAGTGTGGTGGGAGGC 471
             ||||||||||| ||||||||||||| ||  |||||||||||||||  |||||| ||||
Db    137 GlyAlaIleValGlyAlaGlyIleLeuTyrLeuValThrProSerValValGlyGly 156

Qy    472 CCTCCAAAGGGCTGACTCAGCTTTCTGGAAAGCAGTCACAGCGGAATTCTGGCC 36
             |||||||:::|||| |||||||||||||||||:::::||
Db     17 ProProAlaArgSerSerLysGlyAlaArgLeuCysLysGluSerIleGluMetVal 36

Qy    112 GCTTTCAAAGGGGTCTGACTCAAGCTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCC 171
             |||||||| |||||| ||||:::| |||||||||||||||||||||||||:::||
Db     37 AlaPheLysGlyValTrpProPheHisProPheProTrpProTrpProTrpLysAlaValSerAlaGluPheLeuVal 56

Qy    172 ATGCTTATTTTTGTTGTCTCCTCCAGCCTGGGATCCACCATCAACTGGGGTGGAACAGAAAAG 231
             |||||||||||| |||| ||||||:::||||||||||||||||:::||||||:|||||
Db     57 MetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlySerArgSerAsp 196

Qy    232 CCTTTACCGGTCGACATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATG 291
             |||||||||||||   |||||| ||||||  ||||||  ||||||||||||||||
Db     77 ProLeuProValAspMetValLeuIleSerLeuSerPheGlyLeuSerLeuSerIleAlaThrMet 96

Qy    292 GTGCAGTGCTTTGGCCATATCAGCGGGTGGCCATCATCAACCCTGCAGTGACTGTGGCCATG 351
             |||||||||| |||||||||||||||||| ||| ||:::|||||||  ||||||||||
Db     97 ValGlnCysPheGlyHisIleSerGlyValProSerValValProAlaValThrValAlaMet 116

Qy    352 GTGTGCACCAGGAGAGATCAGCATCGCCAAGTGTTCTTCTACATCGACGCCCAGTGCCTG 411
             |||||||||||||| |||||||  ||||| |||:::|||||||||||||||||||
Db    117 ValCysThrArgLysIleSerLeuAlaLysSerLeuAlaLysPheTyrIleLeuAlaGlnCysLeu 136

Qy    412 GGGGCCATCATTGGAGCAGGAATCCTCTATCTCGTCACACCTCCCAGTGTGGTGGGAGGC 471
             ||||||||||| ||||||||||||| ||  |||||||||||||||  |||||| ||||
Db    137 GlyAlaIleValGlyAlaGlyIleLeuTyrLeuValThrProSerValValGlyGly 156
```

Qy 472 CTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTCCTGGTTGAGTGTG 531
              ||||||| ||||||||| ||||| :|||||||||||||| |||||:|||| |:||||
Db 157 LeuGlyValThrAlaValHisGlyAspLeuSerAlaGlyHisGlyLeuLeuValGluLeu 176

Qy 532 ATAATCACATTCAATTGGTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGACTGAT 591
              ||||||||||||||||||||||||||||:||||:|||||||||||||||:|||||
Db 177 IleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgSerAsp 196

Qy 592 GTCACTGGCTCAATAGCTTAGCAATTGCAATTTCGTTGCAATTGGACATTTATTTGCA 651
              |||||||| ||:||||||||||||||||||||||||||||||||||||||
Db 197 ValThrGlySerValAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAla 216

Qy 652 ATCAATTATACTGGTGCCAGCATGAATCCCGCCTTGACCTGCAGTTATCATG 711
              ||||||||||||||| :|||| ||||||||||||||:|||||||||||
Db 217 IleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMet 236

Qy 712 GGAAATTGGAAAACCATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCT 771
              ||| ||||||| |||||||||||||||||||||||||||||||||||||||
Db 237 GlyLysTrpGluAsnGlnIleTrpValTyrTrpValGlyProIleIleGlyAlaValLeuAla 256

Qy 772 GGTGCCCTTTATGAGTATGTCTTCTGTCCAGATGTGAATTCAAACGTCGTTTAAAGAA 831
              |||||| ||||||||||||||| ||||||||||:||||||||||:::
Db 257 GlyAlaLeuTyrGluTyrValTyrCysProAspValGluLeuLysArgArgPheLysAspAsp 276

Qy 832 GCCCTTCAGCAAAGCTGCCCAGCAAACAAAGGAAGCTACATGAGAGGTGAGGACCAACAGG 891
              |||||||:|||| :|||||| :||||||||||||||||:|||||||||||||
Db 277 ValPheSerLysAlaThrGlnProSerLysGlyLysTyrIleGluValAspAspThrArg 296

Qy 892 AGTCAGGTAGAGACGGATGACCTGATTCTCAAAACCTGGAGTGGTGCATGTGATTGACGTT 951
              ||||||||||||||||||| |||||||||||:|||||||||||:||||:|||||
Db 297 SerHisValGluThrAspAspLeuIleLeuLysProGlyIleLeuValHisValIleAspIle 316

Qy 952 GACCCGGGAGGAGGAGGAAGAGGGAAAGACCAATCTGGAGAGGTATTGTCTTCAGTA 1008
              |||||| |||||||||||:||||| ||||||||||||||||||
Db 317 AspArgSerGluAspLysLysGlyLysGlyArgGlyProSerSerGluValLeuSerSerVal 335

RESULT 10
Q8BR89
ID   Q8BR89   PRELIMINARY;   PRT;   224 AA.
AC   Q8BR89;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE   enriched library, clone:B230105H10 product:aquaporin 4, full insert
DE   sequence. (Fragment)
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX   MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]

RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka K., Matsuura J., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR   EMBL; AK045357; BAC32325.1; -.
DR   HSSP; P29972; 1H6I.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019867; C:outer membrane; IEA.
DR   GO; GO:0015288; F:porin activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
KW   Porin; Transmembrane; Transport.
FT   NON_TER   1   1
SQ   SEQUENCE   224 AA;   23938 MW; 6076CD56ACEED029 CRC64;

Alignment Scores:
Pred. No.:   224
Score: 2.46e-75   1103.00
Percent Similarity: 96.88%   214
Best Local Similarity: 95.54%   Matches: 3
Query Match: 50.57%   Conservative: 7
DB: Mismatches: 0
                          Length: Indels: 0
                          Matches: Gaps: 0

US-10-723-180-1 (1-1152) x Q8BR89 (1-224)

Qy 337 GTGACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTACATC 396
              |||||||||||||||||||||||||||||||||||||||||||||||:|||
Db   1 ValThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIle 20

Qy 397 GCAGCCCAGTGCCTGGGGGCCATCATTGGAGCAGGAATCCTCTATCTGGTCACACCTCCC 456
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 IleAlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProPro 40

KW   Porin; Transmembrane; Transport.
SQ   SEQUENCE   218 AA; 23523 MW; AE8AB67A99D4504E CRC64;

Alignment Scores:
Pred. No.:
Score: 1.26e-73
                        1080.50
Percent Similarity: 67.49%
Best Local Similarity: 67.49%
Query Match: 49.54%
DB:

Length: 218
Matches: 218
Conservative: 0
Mismatches: 0
Indels: 105
Gaps: 1

US-10-723-180-1 (1-1152) x Q6L7A0 (1-218)

```
Qy     40 ATGAGTGACAGACCCAGCAAGGCGGTGGGTAAGTGTGGTACCAGAGAG 99
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetSerAspArgProThrArgArgArgTrpGlyGlyLysCysThrArgGlu 20

Qy    100 AACATCATGGTGGCTTCAAAGGGGTCTGGACTCAGCAGTCACAGCG 159
                                     |||  :::
Db     21 Asn-------------------------------------------- 21

Qy    160 GAATTTCTGGCCATGCTCTATTTTGTCCCTCAGCCTGGGATCCACCATCAACTGGGGT 219
Db     21 ---------------------------------------------- 21

Qy    220 GGAACAGAAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGC 279
Db     21 ---------------------------------------------- 21

Qy    280 ATTGCAACCATGGTGCAGTTGCTTTGGCCATATCAGCGGTGGCCACATCAACCCTGCGAGTG 339
Db     21 ---------------------------------------------- 21

Qy    340 ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATGCGCCAAGTCTGTCTTCTACATCGCA 399
Db     21 ---------------------------------------------- 21

Qy    400 GCCCAGTGCCTGGGGCCCATCATTGGAGCAGGAATCCTATCTGGTCACACTTCCCAGT 459
                                         ||||||||||||||||||||||||
Db     22 ----------------IleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer 35

Qy    460 GTGGTGGAGGCCTGGAGTGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCATGGTCTC 519
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     36 ValValGlyIleGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 55

Qy    520 CTGGTTGAGTTGATAATCACATTTCAATTGGTGTGTTACATCTTTGCCAGCTGTGATTCC 579
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     56 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 75

Qy    580 AAACCGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTCTGTTGCAATTGGA 639
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 95

Qy    640 CATTTATTTGCAATCAATATACTGGTGCCAGCATGAATCCGCCGATCCTTTGGACCT 699
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     96 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 115

Qy    700 GCAGTTATCATGGGAAATTGGAAAACCATTGGATATATTGGGTTGGGCCCCATCATAGA 759
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    116 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrThrValGlyProIleIleGly 135

Qy    760 GCTGTCTCCGCTGCTGCTGGGCCTTTATGAGTATGCTCTTCTGTCCAGATGTTGAATTCAAACGT 819
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 AlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGlnPheLysArg 155

Qy    820 CGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAAACAAAGAAGTACATGGAGGTG 879
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnAsnLysGluSerTyrMetGluVal 175

Qy    880 GAGGACAACAGGAGTCAGGTAGACAGCGGATGACCTGAATCTAAAACCTGAGTGTGTGCAT 939
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 GluAspAsnArgSerGlnValGluThrAspSerLeuIleLeuLysProGlyValHis 195
```

Qy   457 AGTGTGGTGGAGGCCTGGGAGTCACCATGGTTCATGGAAAATCTTACCCTGGTCATGGT 516
          || ||||||||
Db   41 SerValValGlyIleGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGly 60

Qy   517 CTCCTGGTTGAGTGATAATCACATTTCAATTGGTGTGTTACATCTTTGCCAGCTGTGAT 576
          |||||||||||||||||||||||||||||||||
Db   61 LeuLeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspAsp 80

Qy   577 TCCAAACGACTGATGTCACTGGCTCAATAGCTTAGCAATTGGATTTCTGTTGCAATT 636
          ||||||||||||||||||||||||||||||||||||
Db   81 SerLysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIle 100

Qy   637 GGACATTTATTTGCAATCAATATACTGGTGCCAGCATGAATCCGCCGATCCTTTGGA 696
          ||||||||||||||||||||||||||||||||||||
Db   101 GlyHisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGly 120

Qy   697 CCTGCAGTTATCATGGGAAATTGGGAAAACCATTGGATATATTGGGTTGGGCCCCATCATA 756
          ||||||||||||||||||||||||||||||||||||
Db   121 ProAlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrThrValGlyProIleMet 140

Qy   757 GGAGCGTGCCTCGCTGGTGGCCTTTATGAGTATGCTCTTCTGTCCAGATGTTGAATTCAAA 816
          |||||||||||||||||||||||||||||||||||:||
Db   141 GlyAlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGlnPheLysLys 160

Qy   817 CGTCGTTTTAAAGAAGCCTTCAGCAAAGCTGCCCAGCAAACAAAGAAGAAGTACATGGAG 876
          ||||||||||||||||||||||||||||||||||||||
Db   161 ArgArgLeuLysGluGluSerLysAlaGlnAlaAlaGlnThrLysGlySerTyrMetGlu 180

Qy   877 GTGGAGGACAACAGGAGTCAGGTAGAGACGGATGACCTTGATTCTAAAACCTGGAGTGGTG 936
          ||||||||||||||||||||||||||||||||||||||
Db   181 ValGluAspAsnArgSerGlnValGluThrGluThrAspAspLeuIleLeuLysProGlyValVal 200

Qy   937 CATGTGATTGACGTTGACCTGGGAGGAGGAGGAGAAGAAGGGGAAAGACCAATCTGGAGAGGTA 996
          |||||||||||||||:::|||||||||||||||||||||||||
Db   201 HisValIleAspIleAspArgGlyGluGluLysLysGlyLysGlyLysAspSerGlyGluVal 220

Qy   997 TTGTCTTCAGTA 1008
          ||||||||||||
Db   221 LeuSerSerVal 224

RESULT 11
Q6L7A0
ID   Q6L7A0;   PRELIMINARY;   PRT;   218 AA.
AC   Q6L7A0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Aquaporin type4 transcript variant c.
GN   Name=AQP4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Ohya S., Imaizumi Y.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR   EMBL; AB128929; BAD22823.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019867; C:outer membrane; IEA.
DR   GO; GO:0015288; F:porin activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000425; MIP.
DR   InterPro; IPR002210; ANK.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.

```
Qy    940  GTGATTGACGTTGACCGGCGGAGAGAGAGAAGAAGGGGAAAGACCAATCTGGAGAGGTATTG  999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196  VallleAspValAspArgGlyGluGluLysGlyLysAspGlnSerGlyGluValLeu      215

Qy   1000  TCTTCAGTA 1008
           |||||||||
Db    216  SerSerVal 218


RESULT 12
Q6AZD2
ID   Q6AZD2;       PRELIMINARY;    PRT;    320 AA.
AC   Q6AZD2;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Zgc:100922 protein.
GN   Name=zgc:100922;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Singapore local strain; TISSUE=Embryo;
RX   PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Singapore local strain; TISSUE=Embryo;
RA   Director MGC Project;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR   EMBL; BC078213; AAH78213.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Transmembrane; Transport.
SQ   SEQUENCE 320 AA; 34427 MW; 6C9B4986C5A17C19 CRC64;


Alignment Scores:
Pred. No.:              2.44e-70          Length:        320
Score:                  1037.50           Matches:       199
Percent Similarity:     77.71%            Conservative:  45
Best Local Similarity:  63.38%            Mismatches:    63
Query Match:            47.57%            Indels:        7
DB:                                       Gaps:          4
```

```
US-10-723-180-1 (1-1152) x Q6AZD2 (1-320)

Qy    73  AAGTGTGGACCTTTGTGTACCAGAGAGAAC---ATCATGGTGGCTTCAAAGGGGTCTGG  129
          ::||:||           |||::::         ||||| |||||||||||||||||||
Db    12  ArgCysValSerSerCysSerCysAsnSerIleMetAlaAlaPheLysGlyValTrp     31

Qy   130  ACTCAAGCTTTCTGAAAGCAGTCACAGCCGGAATTTCTGGCCATGTTATTTTGTTCTC  189
          ||||||||||||||||||||||||||||||||::||||||||||||||:||||||||
Db    32  ThrGlnGluPheTrpArgAlaValSerGlyAlaMetIleIleLeuPheValLeu      51

Qy   190  CTCAGCCTGGGATCCACCATCAACTGGGGTGA---ACAGAAAAGCCTTTACCGGTCGAC  246
          |||||||||||||||||||||||||||||||     ||||||||||||||:::||||||
Db    52  LeuSerLeuGlySerThrIleAsnTrpGlyAlaLysGlnProProProAlaAsp      71

Qy   247  ATGGTTCTCATCTCCCTTGCTTGGACTCAGCATTGCAACCATGGTGCAGTGCTTTGGC  306
          ::|||||||||||||||||||||||||||||:|||:::||||||||||||||||||||
Db    72  LeuValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrLeuValGlnCysPheGly  91

Qy   307  CATATCAGCGGTGGCCACATCAACCCTGCAGTGACTGTGGCCATGGTGTGCACCAGGAAG  366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    92  HisIleSerGlyAlaHisIleAsnProAlaValThrValAlaMetValAlaThrArgLys  111

Qy   367  ATCAGGCATCGCCAAGTCTGTCTTCACATCGCCAGCCCAGTGCCTGGGGCCATCATTGGA  426
          ::::||||:|||:|||||                |||||||||::|||||||||||||||
Db   112  LeuSerLeuAlaLysGlyValPheTyrLeuLeuAlaGlnCysLeuGlyAlaValValGly  131

Qy   427  GCAGGAATCCTCATCTGGTCACACCTCCCAGTGTGGTGGGAGGGCCTGGAGTCACCATG  486
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132  AlaAlaIleLeuTyrTrpGlyValThrProAlaSerValArgGlyMetGlyValThrSer  151

Qy   487  GTTCATGGAAATCTTACCGCTGGTCATGGTCTCCTGGTTGAGTTGATAATCACATTTCAA  546
          :::::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   152  ValAsnGluGluIleSerAlaGlyHisHisAlaIleValIleGluLeuIleLeuThrPheGlu  171

Qy   547  TTGGTGTTTACTATCTTGCCAGCTGTGATTCCAAACGACTGTACTGGCTCAATA  606
          ||||||||||||||||||||||||||||||||||||||||||  |||||||:|
Db   172  LeuValPheThrValPheAlaThrCysAspProLysArgAsnAspLysGlySerAla  191

Qy   607  GCTTTAGCAATTGGATTTCTGTTGCAATTGACATTATTTGCAATCAATTATACTGGT  666
          ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db   192  AlaLeuAlaIleGlyLeuSerValCysIleGlyPheAlaIleProTyrThrGly    211

Qy   667  GCCAGCATGAATCCCGCCCATCCTTTGGACCTGCAGTTATCATCGGGAAATTGGAAAAC  726
          ||||||||||||||||||  ||||::::||||||||||||||||||||||||||::::|
Db   212  AlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetValLysTrpGlnAsp  231

Qy   727  CATTGGATATATTGGGTTGGGCCCATCATAGGAGCTGTCCTCGCTGGTGGCCCTTTATGAG  786
          |||||||||||||||||||||||||||||||||||||||  :::||||||||||||||||
Db   232  HisTrpValTyrTrpValGlyProLeuIleGlyIleLeuAlaAlaAlaAlaValTyrGlu  251

Qy   787  TATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCCTTCAGCAAAGCT  846
          ||||:::|||||||||||  :::  ||||||||||:::   |||||||:::|||||||||
Db   252  TyrLeuPheCysProAspAspLeuLysProLeuLysArgArgTyrAlaAspValLeuSerLysSer  271

Qy   847  GCCCAGCAAACAAAAGGAGGAGCTACATGGAGGGTGGAGGACAACAGGAGTCAGGTAGACACG  906
          ||||||||||||||||||||::: :::||||||||||||::: ||:::||||||||||||||
Db   272  ProPheGlnMetGlu---ProTyrArgValValAspThrAspSerTyrProSerAspGln  290

Qy   907  GATGACCTGATTCTAAAACCTGGAGTGGTCATGTGATTGACGTTGACCGGGGAGAGGAG  966
          |||||||||||::: ||    |||||||||||||||||||||||||||||||||||||
Db   291  AlaGlnLeuMetAlaLysGlnAlaAlaLeuAlaLeuArgValLeuAspLeu------Glu  306

Qy   967  AAGAAGGGGAAAGACCAATCTGGAGAGGTATTGTCTTCAGTA 1008
          ||||||     :::::||   :::||||||||||||||||||
Db   307  LysLysGlyArgGluSerThrGlyValValLeuSerSerVal  320
```

RESULT 13
O46426
```
ID   O46426;       PRELIMINARY;    PRT;    178 AA.
AC   O46426;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
```

DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Aquaporin 4 (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand;
RA   Carter E.P., Umenishi F., Matthay M.A., Verkman A.S.;
RT   "Increased water permeability across the blood-gas barrier in rabbit
RT   lungs in the first 24 hours after birth.";
RL   J. Clin. Invest. 0:0-0(1997).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR   EMBL; AF000312; AAB94409.1; -.
DR   HSSP; P29972; 1H6I.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019867; C:outer membrane; IEA.
DR   GO; GO:0015288; F:porin activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
KW   Porin; Transmembrane; Transport.
FT   NON_TER   1   1
FT   NON_TER   178   178
SQ   SEQUENCE   178 AA; 19236 MW; CE43F9E18ACBAEA2 CRC64;

Alignment Scores:
Pred. No.:
Score: 1.09e-58
Percent Similarity: 884.00
Best Local Similarity: 96.63%
Query Match: 94.38%
DB: 40.53%

Length: 178
Matches: 168
Conservative: 4
Mismatches: 6
Indels: 0
Gaps: 0

US-10-723-180-1 (1-1152) x O46426 (1-178)

```
Qy   454 CCCAGTGTCGTGGGAGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGGTCAT 513
             |||| ||||| || |||| |||||||||||||||||||||||| |||:|||||||||||
Db     1 ProGlyValValGlyGlyLeuGlyLeuValThrThrValHisGlyAsnLeuSerAlaGlyHis  20

Qy   514 GGTCTCCTGTTGAGTTGAATAACACATTCAATTGGTGTTACTATCTTTGCCAGCTGT 573
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 GlyLeuLeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheProSerCys  40

Qy   574 GATTCCAAACGGACTGATGTCACTGGCTCAATAGCTTTAGCAATTGGATTTTCTGTTGCA 633
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 AspSerLysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAla  60

Qy   634 ATTGGACATTTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCCGCCCGATCCTTT 693
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IleGlyHisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPhe  80

Qy   694 GGACCTGCAGTTATCATGGGAAATTGGGAAAACCATTGGATATATTGGGTTGGGCCATC 753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 GlyProAlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIle 100

Qy   754 ATAGGAGCTGTCCTCGCTGGTGGCCTTTATGAGTGTCTTCTGTCCAGATGTTGAATTC 813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 IleGlyAlaValLeuAlaGlyTyrGlyLeuTyrGluTyrGluValPheCysProAlaLeu 120

Qy   814 AAACGTCGTTTAAAGAAGCCTTCAGCAAAGAAGCTTGCCCAGCAAACAAAAGAAGCTACATG 873
             :::|||||||||||||||||||||||||||||||||||||||||||||||||:::||
Db   121 LysArgArgPheLysLysProSerAlaLysLysLeuAlaGlnGlnThrLysGluLysSerTyrMet 140

Qy   874 GAGGTGGAGGAGCAACAGGAGCTCAGGTAGAGACGGATGACCTGATTCTAAAACCTGGAGTG 933
             |||||:||||||||||||||||||||||||||||||||:|||:::|||||||||||||||
Db   141 GluValGluAspAsnArgArgSerGlnValGluThrGlyAspLeuIleLeuLeuLysProGlyVal 160
```

Db   934 GTTCATGGTGATTGACGTTGACCGGGAGAGGAGAGAAGAAGGGAAAGACCAATCT 987
             :::||||||||||||:::|||||:||||||||||||||||||||||||||||||
Qy   161 MetHisValIleAspIleAspArgArgGluArgArgLysLysGlyLysAspProSer 178

RESULT 14
Q8K4M1
ID   Q8K4M1   PRELIMINARY;   PRT;   149 AA.
AC   Q8K4M1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Aquaporin-4 isoform M1 (Fragment).
GN   Name=Aqp4;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20419017; PubMed=10960499;
RA   Zelenin S., Gunnarson E., Alikina T., Bondar A., Aperia A.;
RT   "Identification of a new form of AQP4 mRNA that is developmentally
RT   expressed in mouse brain.";
RL   Pediatr. Res. 48:335-339(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bondar A., Alikina T., Zelenin S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Alikina T.Y., Zelenin S.M., Bondar A.A.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR   EMBL; AF219992; AAM81576.1; -.
DR   HSSP; P29972; 1H6I.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019867; C:outer membrane; IEA.
DR   GO; GO:0015288; F:porin activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Porin; Transmembrane; Transport.
FT   NON_TER   1   1
FT   NON_TER   149   149
SQ   SEQUENCE   149 AA; 15663 MW; BA0E8F89359B708C CRC64;

Alignment Scores:
Pred. No.:
Score: 3.08e-45
Percent Similarity: 707.00
Best Local Similarity: 94.59%
Query Match: 91.22%
DB: 32.42%

Length: 149
Matches: 135
Conservative: 5
Mismatches: 8
Indels: 0
Gaps: 0

US-10-723-180-1 (1-1152) x Q8K4M1 (1-149)

```
Qy    40 ATGAGTGACAGACCCAGCAAGGCGGTGGGTAAGTGTGGACCTTTGTGTACCAGAGAG  99
             |||||||||||||||||||||||||||||||||||||||||||||| |||:|||||
Db     1 MetSerAspArgAlaLeuAlaAlaArgTrpGlyHisSerCysSerArgGlu  20

Qy   100 AACATCATGGTGGCTTCAAAGGGGTCTGACTCCAGCTTTCTGAAAGCAGTCACAGCG 159
             :::||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db    21 SerIleMetValAlaSerLysGlyValTrpArgGlyValTrpThrGlnAlaPheThrLysAlaValSerAla  40

Qy   160 GAATTTCTGGCCATGCTATTTTGTTCTCCTCAGCCTGGGATCCAACTCAACTGGGGT 219
             |||||||||||||||||||||||||||:::||||||||||:||:::||||||||||
Db    41 GluPheLeuAlaThrLeuPheLeuLeuLeuGlyValGlySerThrIleAsnTrpGly  60
```

```
FT   TRANSMEM       101    121                3 (Potential).
FT   DOMAIN         122    138                4 Extracellular (Potential).
FT   TRANSMEM       139    159                4 (Potential).
FT   DOMAIN         160    169                5 Cytoplasmic (Potential).
FT   TRANSMEM       170    190                5 (Potential).
FT   DOMAIN         191    211                6 Extracellular (Potential).
FT   TRANSMEM       212    232                6 (Potential).
FT   DOMAIN         233    249                6 Cytoplasmic (Potential).
FT   SITE            82     84                NPA 1.
FT   SITE           198    200                NPA 2.
SQ   SEQUENCE       249 AA;  26111 MW;  F8584962C5F169FE CRC64;
```

Alignment Scores:
```
Pred. No.:                 1
Score:                     2.24e-41
Percent Similarity:        69.72%
Best Local Similarity:     54.13%
Query Match:               30.10%
DB:                        1
```

```
Length:            249
Matches:           656.50
Conservative:      118
Mismatches:        34
Indels:            63
Gaps:              3
                   1
```

US-10-723-180-1 (1-1152) x AQP_AEDAE (1-249)

```
Qy   142  TGGAAAGCAGTCACAGCGGAATTCTGGCCATGCTTATTTTGTCTCCTCAGCCTGGGA  201
               |||  ::: :::      |||  :|| |||:     ||| ::|:   ::|:|||
Db    23  TrpArgMetLeuValAlaGluPheLeuGlyThrPheLeuValSerIleGlyIleGly   42

Qy   202  TCCACCACTCAACTCGGGTGGAACAGAGAAAGCCTTTACCGGTCGACATGGTTCTCATCTCC  261
               |||:||||||  ||||||||                       :||::|||||
Db    43  SerThrMetGlyTrpGlyGlyAspTyrAlaPro-------ThrMetThrGlnIleAla   59

Qy   262  CTTTGCTTTGGACTCAGCATTGCAACCATGGTGCAGTGTGTTGGCCATATCAGCGGTGGC  321
               |||||||         |||   :::|||||||||    ||| |||||::|||||
Db    60  PheThrPheGlyLeuValAlaThrLeuAlaGlnAlaPheGlyHisValSerGlyCys   79

Qy   322  CACATCAACCCTGCAGTGACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAG  381
               |||||||||:: :      ||| ||||||||:    :::::| ::
Db    80  HisIleAsnProAlaValThrIleGlyLeuMetIleThrAlaAspIleSerIleLeuLys   99

Qy   382  TCTGTGTTCTCATCGCCAGTGCCCAGTGCCCATCATTGGAGCAGGAAATCCTCTAT  441
               ||||||||||  ::||||||      |||||||:|||||||    :::::
Db   100  GlyAlaPheTyrIleValSerGlnCysValGlyAlaIleAlaGlyAlaLeuIleLys  119

Qy   442  CTGGTCACACTCCCAGTGTGGTGGGAGGCCTGGGAGTCACCATGGTTCATGGAAATCTT  501
               |||||           |||:||||||||||    :::
Db   120  AlaAlaThrProSerAspValIleGlyGlyLeuGlyValThrGlyValIleAspProArgLeu  139

Qy   502  ACCGCTGGTCATGGTCTCCTGGTTGAGTTGATATTCATTGGTGTTTACTATC  561
               |||||||||       |||||||:||:    |||::|| :::
Db   140  ThrAlaGlyHisGlyValMetGluAlaLeuIleThrPheIleLeuValPheValVal  159

Qy   562  TTTGGCCAGCTGTGATTCCAAACGACGATGATGTCACTGCTGCCAAATAGCTTTAGCAATTGGA  621
               |||:||||||      |||:||||:: |||:
Db   160  HisGlyValCysAspAsnArgSerAspIleLysGlySerAlaProLeuAlaIleGly  179

Qy   622  TTTTCGTTGCAATTGGACATTATTTGCAATCAATTATACTGGTGCCAGCATGAATCCC  681
               ||||           |||||:|||||| :::
Db   180  LeuSerIleThrAlaGlyHisLeuSerAlaIleLeuTyrThrGlyAlaSerMetAsnPro  199

Qy   682  GCCCGATCCTTTGACCTGCAGTTATCATGGGAAATTGGAAAAACCATTGGATATATTGG  741
               |||||||||       |||||||||||||:
Db   200  AlaArgSerPheGlyProAlaValValMetGlyAsnTrpThrAspGlnTrpTyrTrp  219

Qy   742  GTTGGGCCCATCATAGGAGCTGTCCTCCGCTGGCCGCCTTTATGAGTAGTGTCTTC  795
               ||||||||      :::||||||  :::||||||
Db   220  ValGlyProIleValGlyIleLeuAlaGlyIleAlaValTyrArgLeuPhePhe  237
```

Search completed: September 8, 2005, 12:13:42
Job time : 238 secs

```
Qy   220  GGAACAGAAAGCCTTTACCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGC  279
               |||  :::||   |||  ||||||||||          ||||||||||||||
Db    61  GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer   80

Qy   280  ATTGCAACCATGGTGCAGTGTGTTGGCCATATCAGCGGTGGCCACATCAACCCTGCAGTG  339
               |||||||||||||||||||||||| ||||||||:||||||  |||||||:||:||||
Db    81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal  100

Qy   340  ACTGTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCCAAGTCTGTCTTCTCACATCGCA  399
               |||||||||||||||     |||  |||:|||||||| ||| ||||| :||:|||
Db   101  ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleIle  120

Qy   400  GCCCAGTGCCTGGGGCCCATCATTGGAGCAGGAAATCCTCTATCGGTCACACTCCCAGT  459
               |||||||| |||||||||||:||||||||||||  |||||| |:||||||||||
Db   121  AlaGlnCysLeuGlyLeuAlaIleIleGlyAlaGlyAsnProLeuLeuTyrLeuValThrProProSer  140

Qy   460  GTGGTGGGAGGCCTGGGAGTCACC  483
               ||||||||||||
Db   141  ValValGlyGlyLeuGlyValThr  148
```

RESULT 15
```
AQP_AEDAE
ID   AQP_AEDAE      STANDARD;       PRT;     249 AA.
AC   Q9NHW7;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Aquaporin AQPAe.a.
OS   Aedes aegypti (Yellowfever mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX   NCBI_TaxID=7159;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UTMB; TISSUE=Malpighian tubules;
RX   MEDLINE=20428297; PubMed=10971718;
RA   Pietrantonio P.V., Jagge C., Keeley L.L., Ross L.S.;
RT   "Cloning of an aquaporin-like cDNA and in situ hybridization in adults
RT   of the mosquito Aedes aegypti (Diptera: Culicidae).";
RL   Insect Mol. Biol. 9:407-418(2000).
CC   -!- FUNCTION: Forms a water-specific channel (By similarity).
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Localized in tracheolar cells associated with
CC       the Malpighian tubules.
CC   -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC       three membrane-spanning domains and a pore-forming loop with the
CC       signature motif Asn-Pro-Ala (NPA).
CC   -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; AF218314; AAF64037.1; -.
DR   HSSP; P29972; 1H6I.
DR   InterPro; IPR000425; MIP.
DR   Pfam; PF00230; MIP; 1.
DR   PRINTS; PR00783; MINTRINSICP.
DR   ProDom; PD000295; MIP; 1.
DR   TIGRFAMs; TIGR00861; MIP; 1.
DR   PROSITE; PS00221; MIP; 1.
KW   Repeat; Transmembrane; Transport.
FT   DOMAIN           1     26                1 Cytoplasmic (Potential).
FT   TRANSMEM        27     47                1 (Potential).
FT   DOMAIN          48     57                2 Extracellular (Potential).
FT   TRANSMEM        58     78                2 (Potential).
FT   DOMAIN          79    100                3 Cytoplasmic (Potential).
```

This Page Blank (uspto)